THE ROLE OF THE LONG TERMINAL REPEAT (LTR) IN THE PATHOGENESIS OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 2 (HIV-2)

BY

EMMA LOUISE BAKER

Department of Virology,
University College London Medical School,
London,
United Kingdom,
W1P 6DB.

This thesis is submitted to the University of London
in fulfilment for the degree of
Doctor of Philosophy (Ph.D)
2004.

The work described within this thesis has been produced
by E.L.Baker unless otherwise stated.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DEDICATION

This thesis is dedicated to my partner Tim Chamberlin,
whose love and complete belief enabled me
to carry on until the last,
and to my parents,
for their inspiration and continuing love and support.

ACKNOWLEDGEMENTS

I would like to thank all my colleagues and collaborators within The Department of Virology at University College London Medical School, and The Medical Research Council Laboratories in Fajara, The Gambia. In particular, I wish to thank my two supervisors Prof. Richard Tedder and Dr. Neil Berry for their encouragement and guidance throughout my studies, and Drs Hilton Whittle and Koya Ariyoshi without whom these studies would not have been possible.

I would like to acknowledge the input of Dr. Peter Balfe in the design and supervision of this work. I am also grateful to Dr Judith Breuer at The Royal London School of Medicine and Dentistry, and Dr David Markovitz at the University of Michigan Medical Centre for their support and advice.

Finally, I would like to thank all my fellow Ph.D friends at UCLMS in particular Samreen, Phil, Paul, Kev, Dave and Anna for the support they gave and the fun we had despite the tough times, and to Hilda for her much needed company and continuing words of encouragement.

This degree was funded by a Medical Research Council Ph.D studentship. The University College Graduate School Presidents Fund kindly donated additional funds for travel to conferences in The Gambia and Colorado, USA. In addition, a six-week visit to Dr David Markovitz's laboratory at the University of Michigan Medical Centre was funded by a Bogue Fellowship.

ABSTRACT

Recent data has indicated that the prolonged asymptomatic phase and freedom from clinical illness experienced by the majority of HIV-2 infected individuals results from a lower level of virus production during infection than that observed in either HIV-1 infection or in the small number of HIV-2 infected individuals displaying relatively rapid progression to disease. The rate at which viral gene products are transcribed is one mechanism by which the rate of virus production can be controlled and is mediated by the Long Terminal Repeat (LTR) regions of the HIV-2 genome.

A limiting dilution sensitive nested PCR has been developed to amplify the HIV-2 LTR from clinically and phenotypically characterised infected sources, including a number of HIV-2 isolates and uncultured PBMC samples derived from Gambian HIV-2 infected long-term non-progressor (LTNP) and rapid progressor (RP) patients. Using a highly efficient cloning system LTR amplicons have been cloned into a firefly luciferase reporter vector. A dual luciferase reporter assay has been used to determine LTR-directed basal and Tat *trans*activated levels of transcription in physiologically relevant cell lines. LTRs derived from LTNP patients tended to direct lower levels of basal and Tat *trans*activated transcription when compared to LTRs derived from RP patients. The difference between the two groups was more pronounced in the T cell-like Jurkat cell-line. Nucleotide sequence analyses of the LTR clones has revealed that the rate of general and G-to-A mutations at single functional sites within the LTR is higher in sequences derived from LTNP patients compared to RP patients. Taken together our data implies a relationship between LTR activity, sequence variation, and overall levels of virus expression in HIV-2 as evidenced by higher levels of circulating peripheral HIV-2 RNA in the patients with progressive disease profiles. Therefore, different LTR responsiveness may relate to different rates of virus production and disease progression rates and hence be a determinant in viral pathogenesis.

COMMON ABBREVIATIONS

A	adenine
AIDS	Aquired Immune deficiency Syndrome
Amp	ampicillin
bp	base pair
C	cytosine
DNA	deoxy-ribonucleic acid
dNTP	2' deoxyribonucleotide 5' triphosphate
ds	double stranded
E.coli	Eschericia coli
EDTA	ethylene-diaminetetra-acetic acid
G	guanine
g, mg, μg, ng, pg	gram, milligram, microgram, nanogram, picogram
g	acceleration due to gravity
gp120/gp105	HIV-1 120kD/HIV-2 105kD Surface envelope glycoprotein
gp41/gp36	HIV-1 41kD/HIV-2 36kD Transmembrane envelope glycoprotein
HIV	Human Immunodeficiency Virus
Kb	kilobasepair(s) (K=1000)
kD	kiloDalton
LTR	Long Terminal Repeat
M, mM, μM	molar, millimolar, micromolar
ml, mls, μl	millilitre, millilitres, microlitre
MRC	Medical Research Council
mRNA	messenger RNA
oligo	oligonucleotide
PBS	phosphate buffered saline
PCR	polymerase chain reaction
RNA	ribonucleic acid
SIV	Simian Immunodeficiency Virus
T	thymidine
TAE	Tris-acetate EDTA buffer
TAR	Transactivation Responsive Region
TE	Tris-HCL EDTA
Tris	tris (hydroxymethyl) aminomethane
U	uracil
V	volts

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1.
General Introduction.

1.1 The Retroviridae

To aptly describe the remarkable nature of the retroviridae is an extremely difficult undertaking. The wide-ranging appeal to study this distinct family of RNA viruses surely stemmed from the discovery that retroviruses unequivocally contradicted the central dogma of molecular biology. In 1964 retrovirologist Howard Temin illustrated that infection with RNA tumour viruses such as avian sarcoma virus, was blocked using inhibitors of DNA synthesis. This finding suggested that DNA synthesis was required for the multiplication of RNA tumor viruses, indicating that genetic information could flow from RNA to DNA. From this, Temin went on to propose that the information of the infecting RNA was transferred to a DNA copy, a provirus, which then served as a template for the synthesis of viral RNA. The discovery of the RNA dependent DNA polymerase enzyme, reverse transcriptase, made independently by Temin (Temin & Mizutani *et al.,* 1970) and Baltimore (Baltimore, 1970) in 1970, verified Temin's intrepid hypothesis. The subsequent discovery that all RNA tumour viruses encoded a reverse transcriptase prompted the family of viruses to be termed retroviruses.

In addition to the contradictory flow of genetic information directed by members of the retroviridae, retroviruses exhibit an amazing capacity to intimately associate with host DNA. Other animal and plant viruses have been found to produce provirus using reverse transcriptase as part of their replication cycle, but retroviruses are the only viruses possessing a mechanism that enables stable proviral integration into host DNA. The effects caused upon integration of the retroviral provirus go someway toward determining the diverse pathogenesis observed within the retrovirus family, and can be used to distinguish between its different members.

The strikingly intimate relationship the retrovirus shares with its host is exemplified by the ability of the retroviral genome to be synthesized and processed entirely by host cell transcriptional machinery. As a result, all retroviral genomes contain a 5' capping group and a 3' polyadenylation signal; features that are normally only found in newly made eukaryotic mRNA transcripts. In addition, the retroviral genome is associated with a highly conserved specific tRNA enabling it to prime genome replication. Taken together, the retrovirus family exhibits the most desirable attributes to aid complete integration into host cell processes, enabling successful virus host relationships to be established during viral infection.

Retroviruses are associated with wide ranging diseases, primarily of vertebrates, including immunodeficiencies, neurological disorders, and malignancies which can be rapid or of long latency. The majority of the pathological conditions caused by retroviruses are as a result of infection with exogenous retroviral elements, however, retroviruses may also be present as endogenous elements. Endogenous retroviruses are not generally associated with pathogenic infection; they exist as integrated provirus within the germ line of many vertebrates and are inherited via vertical transmission. The endogenous proviruses are usually transcriptionally silent and are often defective, although some can be induced to express and replicate at low levels. The existence of such an established but largely non-pathogenic virus-host relationship highlights the vast diversity of virus - host interactions characteristic of the retrovirus family.

1.1.1. Retrovirus morphology and general genomic organisation

Despite the variety of pathological conditions relating to retrovirus infection there appears to be a basic structural and genetic pattern displayed by each member of the retrovirus family. Generally, mature virions occur as enveloped particles 100nm in diameter. Retroviruses have a diploid genome consisting of two molecules of positive sense single stranded RNA. The virion genome is encapsulated within a spherical or conical shaped internal nucleocapsid. As with virion structure, genomic organisation is conserved within the retrovirus family. All retrovirus genomes contain three structural genes that are identically orientated 5' to 3' within the genome. The group-specific antigen (*gag*) gene encodes the virion core proteins, the polymerase (*pol*) gene encodes reverse transcriptase and other viral enzymes, and the envelope (*env*) gene encodes the oligomeric glycoprotein spikes that are located in the surrounding host derived lipid envelope. The viral genome is flanked by two Long Terminal Repeat (LTR) sequences that contain viral promoter and enhancer elements that interact with *cis*- and *trans*- acting DNA binding proteins to direct viral gene expression. The basic genetic organisation of the retrovirus genome is always LTR-*gag-pol-env*-LTR.

Thin-section electron micrograph studies have revealed that despite the commonality of the virion structure (Gelderblom *et al.*, 1972), retroviruses can be classified into four morphological groups (types A to D) in addition to the previously described genera classifications. A-particles exist as roughly spherical double-walled intracellular structures, 60-90nm in diameter. They are non-infectious, produced by a number of endogenous proviruses, and represent the immature precursors of B-type and D-type viruses. In comparison, B- and D- particles represent extracellular virion structures; they exist as encapsulated versions of A-particles. The B-particle envelope is characterised by the presence of surface projections, the D-particle envelope does exhibit surface projections but they are less prominent than those found on B particles. C-particles, unlike A-particles, lack visible intracellular structures. Immature core formation occurs simultaneously with virion budding. Maturation of the particle is defined by condensation of the core structure, this occurs within the enveloped particle, a similar process is also observed within B-particles. The envelope of C-particle virions does not exhibit surface spikes.

1.1.2. Retrovirus genera taxonomy

Traditionally, retrovirus taxonomy distinguished members of the retrovirus family by pathogenicity, this lead to the division of three subfamilies (Teich, 1985). The first, the *Oncoviridae,* represented a subfamily of tumorogenic retroviruses that were largely exogenous; with the majority of members associated with pathogenic conditions. The second, the *Spumaviridae,* were classified as exogenous retroviruses that caused persistent but benign infection, characterised by vacuolation of cells in culture. The third, the *Lentiviridae,* were classified as a subfamily of retroviruses that gave rise to slowly developing diseases of the immune system characterised by long incubation periods. The majority of members within this subfamily caused benign infection in their natural hosts, but a minority were associated with fatal pathogenic conditions.

Retroviral classification has been more recently redefined based upon genome structure and nucleotide sequence relationship as well as pathogenicity, this has resulted in the recognition of seven distinct virus genera (Coffin, 1992) (Table 1.1, pg 30) as opposed to the traditional three subfamilies. Only three of the seven genera contain human retroviruses, the HTLV-BLV group, the Lentivirus group and the Spumavirus group

Previously characterised as the *Oncoviridae* the oncogenic members of this retrovirus family are now classified into five distinct genera (Table 1.1, pg 30). Four of the five groups contain retroviruses that infect avian and mammalian species; the remaining group has been shown to cause pathogenic conditions in humans. Studies investigating the structural and functional relationship between viral oncogenes and cellular proto-oncogenes have centred on retroviral examples from all five genera. Information provided by such studies have greatly contributed to our understanding of the molecular mechanisms involved in retroviral induced malignancy, signal transduction and cellular growth.

Oncogenic retroviruses have the ability to induce transformation and tumour formation by controlling the expression of viral or cellular oncogenes. This control is dictated by the ability of the retrovirus to either integrate into host DNA or incorporate parts of cellular proto-oncogenes into their genomes. In either mechanism oncogenic expression is controlled by the viral LTR, hence the cellular proto-oncogene promoter no longer controls expression levels, growth factor regulation or cell specificity. Examples of both mechanisms are exhibited in infection with retroviruses from the five groups discussed below.

The avian leukosis-sarcoma virus (ALSV) group contains exogenous and endogenous retroviruses associated with tumor induction and leukaemias in avian species (Table 1.1, pg 30) (Payne, 1992). The exogenous members of the genera, with the exception of Rous-associated virus (RAV) and avian leukosis virus (ALV) contain oncogenes. Rous sarcoma virus (RSV) contains oncogene v-*src*; this gene is a processed version of the cellular proto-oncogene c-*src*, a tyrosine specific protein kinase, but contains major deletions (Tanaka *et al.*, 1990, Zhou *et al.*, 1990). Similarly, avian erythroblastosis virus (AEV) contains v-*erb*-B a deletion containing counterpart to cellular *erb*-B (Bruskin *et al.*, 1990, Miller *et al.*, 1990). Infection with either of these viruses results in LTR directed expression of oncogenes whose cellular relatives are involved in signal transduction. The changes to the virally expressed oncogenic proteins allow them to behave independently of normal cellular transduction signals and results in the transformation of a normal cell to a malignant one (Miller *et al.*, 1990, Zhou *et al.*, 1990).

Retroviruses that lack oncogenes are able to induce malignancies by their ability to integrate into the host genome and cause a process called insertional activation. Tumours induced by oncogene-lacking retroviruses such as avian leukosis virus (ALV) exhibit a common feature; the provirus is always integrated within a particular gene of the cellular genome. In the case of ALV the proto-oncogene c-*myc* is targeted, upon proviral integration c-*myc* expression is solely controlled by the 3' LTR, and results in the initiation of tumorigenesis (Robinson & Gagnon, 1986, Goodenow & Hayward, 1987).

The mammalian C-type group of retroviruses also contains exogenous and endogenous aetiological agents (Table 1.1, pg 30). The exogenous retroviruses of this genera have been shown to cause leukaemia's and sarcoma's within numerous mammalian species and many are associated with oncogenes (Kozak & Ruscetti, 1992). A distinguishing feature among members of this group appears to be the involvement of virion proteins in the oncogenic process. In murine leukaemia virus (MLV) and feline leukaemia virus (FLV) infection, the *gag* protein associates with v-*abl* and *fms* respectively, ultimately resulting in the activation of T-cell lymphomas (Witte *et al.*, 1978, Hampe *et al.*, 1984, Neil *et al.*, 1987, Neil *et al.*, 1991). The MLV *env* proteins have also been linked with oncogenic induction through stimulation of cellular IL-2 receptors (Li & Baltimore, 1991).

Mouse mammary tumor virus (MMTV) is the best-studied member of the *B-type virus group* and represents the only infectious retrovirus of this genera (Table 1.1, pg 30). Unlike the majority of retroviruses discussed so far MMTV is not associated with any oncogene yet induces sporadic mammary adenocarcinomas in infected female mice (Morris, 1991). Exogenous and endogenous forms of MMTV

exist, although exogenous transmission is also vertical, passing from mother to offspring via ingested milk. Carcinoma induction is stimulated by the random insertion of the proviral genome into a host proto-oncogene (*int* genes). The alteration of expression of *int* genes as a result of the MMTV promoter insertion has been causally linked with proliferation of tumours in the mammary gland (Nusse *et al.*, 1990, Nusse, 1991, Jhappan, *et al.*, 1992).

The *D-Type* group of exogenous and endogenous retroviruses have only been isolated from primate species; the most widely studied example is Mason-Pfizer monkey virus (MPMV) (Table 1.1, pg30). Retroviruses from this genus exhibit A particle morphology, characterised by intracellular core formation prior to virion budding. In this respect, D-type retroviruses are most closely related to B-type viruses and spumaviruses, immature particles from both groups resemble intracisternal A particles. D-type retroviruses containing oncogenes have not been isolated.

Until members of the *Human T-cell lymphotrophic virus/ bovine leukaemia virus* group had been isolated, proof that retroviruses were causally involved in human neoplastic disease was lacking. With the advent of continuously growing cell lines, two groups independently isolated retroviruses exhibiting C type morphology. In 1980, a group lead by B.J. Poiez isolated Human T cell lymphotrophic virus (HTLV-1) from a patient exhibiting cutaneous T-cell lymphoma (Poiez *et al.*, 1980). Two years later Yoshida *et al* isolated a C type retrovirus from a patient suffering from adult T cell leukaemia termed adult T cell leukaemia virus (ATLV) (Yoshida *et al.*, 1982). Following serologic and nucleic acid analysis ATLV was found to exhibit equivalent divergence to isolates of HTLV-1 and was deemed a member of the HTLV group. HTLV-1 has been linked to a number of other diseases distinct from T cell lymphomas/lymphomas including tropical spastic paraparesis, HTLV-associated myelopathy, and immunosuppressive disorders. Unlike the majority of other oncogenic retroviruses, exogenous HTLV-1 is not associated with any oncogene or integration site. HTLV-1 integrates into the host cell genome randomly and detectable expression of viral genes is not observed. It appears that oncogenesis is in some way related to the ability of the integrated HTLV-1 provirus to induce constant T cell proliferation. A malignant clone may subsequently arise and manifest into a tumor, the length of time that this process takes to occur may determine the long latency seen in HTLV infection.

The HTLV genome encodes the three major retroviral genes, *gag, pol,* and *env*, and two additional regulatory genes, *tax* and *rex*. In this respect, HTLV shows similarity to members of the *lentiviridae*, the other genera of the retrovirus family to encode regulatory genes and cause disease in humans. The *tax*

protein *trans*activates transcription directed by the HTLV-1 LTR using an indirect mechanism distinct from but related to *trans*activation of lentivirus transcription. By directly interacting with transcription factors NF-κB, CREB, and AP-1, *tax* is able to up-regulate levels of transcription without having to physically bind to the HTLV enhancer and promoter elements. *Tax* has also been shown to exert upregulatory effects upon cellular and viral promoter regions distinct from those found within the HTLV-1 genome. The ability of *tax* to *trans*activate IL-2 and IL-2R promoter regions has been implicated in the transforming and subsequent oncogenic potential of HTLV-1. In direct analogy with members of the lentiviral group the regulatory protein *rex* regulates HTLV-1 gene expression post-transcriptionally acting via a *rex* responsive element. Transmission of HTLV occurs via sexual intercourse, infected blood products, and breast milk.

A second retrovirus associated with T cell malignancies has been identified that is related to but distinct from HTLV-1. This retrovirus, now termed HTLV-2, was isolated from a patient suffering from atypical hairy-cell leukaemia (Kalyanaraman *et al.,* 1982). HTLV-2 predominantly infects $CD8^+$ cells as opposed to $CD4^+$ cells, the preferred target of HTLV-1, however the causal link between HTLV-2 and leukaemia is less palpable than with HTLV-1. Nevertheless, there have been other reports of HTLV-2 in association with T cell malignancy, and isolation of this virus from a second patient with hairy-cell leukaemia provides further evidence for the etiologic role of HTLV-2 in this disease. A link between HTLV-2 and intravenous drug abusers (IVDA) has been widely reported with higher prevalence of infection being found in American patients in comparison to European patients. Neurodegenerative disorders have also been linked with HTLV-2 infection indicating that diseases found to be associated with HTLV-2 infection may mirror those linked to infection with HTLV-1.

Approximately 5% of HTLV-1 infected individuals develop symptomatic disease; considering the world wide infected population is thought to number between 10 and 20 million there is a significant need for vaccine strategies. Two animal models exist for vaccine development, bovine leukaemia virus (BLV), and simian T cell leukaemia virus type 1 (STLV-1). Subunit vaccine trails using monkeys have indicated that protection against STLV can be achieved however a vaccine candidate for human clinical trails has yet to be produced.

The *Spumavirus group* represent a group of retroviruses more commonly known as foamy viruses (Table 1.1, pg 30). Widely distributed among non-human primates, felines, and bovines, foamy viruses exhibit a

large cellular tropism and are highly lytic *in vitro*, forming giant multinucleated cells (syncitia) containing vacuoles giving them a foamy appearance, hence the term "foamy". Despite this distinguishing feature, foamy viruses appear to be innocuous *in vivo*, evidence linking them to disease within their natural hosts and humans remains tenuous. To date there has only been one reported isolation of human foamy virus (HFV) (Werner & Gledерblom, 1979), the remaining reports of human foamy virus infection have resulted from zoonotic transmission of simian foamy virus (SFV) from natural host non-human primate sources (Heneine *et al.*, 1998). The genome of HFV consists of the three structural genes common to all retroviruses; with a further four open reading frames (ORF's), bel-1, 2, 3 and S1, which encode accessory proteins. *Bel-1* encodes the HFV *trans*activator protein, which shares homology with the *trans*activator protein encoded by members of the Lentiviridae, and represents the only accessory protein that is essential to HFV replication (Rethwilm *et al.*, 1991).

Members of the *Lentivirus group* form a distinct group of exogenous retroviruses, which cause infections associated with prolonged incubation periods prior to disease onset (Table 1.1, pg 30). Lentiviruses are capable of persistently infecting a vast array of cellular targets although they typically infect cells of the immune system, primarily those of monocyte/macrophage lineage.

The first member of the *lentiviridae* was identified in 1904. Equine infectious anaemia virus (EIAV) was found to cause autoimmune haemolytic anaemia and in some instances encephalopathy in horses (Vallee & Carre, 1904). A second disease characterised by slow progression was identified in Icelandic sheep in 1957 (Sigurdsson *et al.*, 1957). The etiologic agent causing the slowly progressive pneumoencephalitic disease in sheep was found to be of retroviral origin and was identified as maedi/visna virus (MMV), the second member of the lentivirus family (Narayan & Clements, 1989). An arthritis/encephalitis virus of goats (caprine arthritis-encephalitis virus, or CAEV) exhibiting different properties from MMV has also been identified and characterised as a lentivirus (Saltarelli *et al.*, 1990).

More recently, the molecular and biological characterisation of retroviruses that cause severe immunodeficiencies have resulted in their addition to the lentiviral family. Among those characterised are the human immunodeficiency viruses type 1 and type 2 (HIV-1, HIV-2), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), and simian immunodeficiency virus (SIV). All have been causally associated with severe immunodeficiencies. The immunodeficiency viruses grouped into

the lentivirus family having a similar genomic organisation and exhibit type C morphology. These are discussed in more detail in Section 1.2.

The Retroviridae:

Genus	Species
ALSV-group	Rous sarcoma virus (RSV) Avian erythroblastosis virus (AEV) Avian myeloblastosis virus (AMV) Rous-associated virus (RAV) RAV-0
Mammalian C-type	Moloney murine leukaemia virus (Mo-MLV) Abelson murine leukaemia virus (A-MuLV) AKR-MuLV Feline leukaemia virus (FeLV) Hamster leukosis virus Simian sarcoma virus (SSV)s Primate endogenous viruses (BaEV), (MAC-1), (OMC-1). Reticuloendotheliosis virus (REV)
B-Type viruses	Mouse mammary tumor virus (MMTV)
D-Type viruses	Mason-Pfizer monkey virus (MPMV) "SAIDS" viruses Simian Retrovirus (SRV)
HTLV-BLV group	Human T-cell leukaemia (or lymphotropic) virus, (HTLV-1, HTLV-2) Bovine leukaemia virus
Lentivirus	Human immunodeficiency virus, (HIV-1, HIV-2) Simian immunodeficiency virus (SIV) Feline immunodeficiency virus (FIV) Bovine immunodeficiency virus (BIV) Visna/maedi virus Caprine arthritis-encephalitis (CAEV) Equine infectious anaemia virus (EIAV)
Spumavirus	Simian foamy virus (SFV) Human foamy virus (HFV) or Human spumareetrovirus (HSRV) Feline syncytium-forming virus (FeSV)

Table 1.1 Classification of the retroviridae (adapted from Davis *et al.*, 1990).

1.2 Lentiviral infections associated with immunodeficiency

1.2.1. Bovine immunodeficiency virus

Bovine immunodeficiency virus (BIV) was originally isolated from a cachectic cow displaying lymphadenopathy and mild encephalitis (Van der Maaten *et al.,* 1972). Since this isolation however, few investigators have reported a pathogenic link between BIV infection and disease with the possible exception of Carpenter *et al* in 1992 (Carpenter *et al.,* 1992). Calves inoculated with BIV were found to develop mild follicular hyperblastia in the lymph nodes, however a similar study carried out by a different group failed to link BIV infection with clinical disease (Onuma *et al.,* 1992). The significance of this pathogen in its natural host remains undetermined.

1.2.2 Feline immunodeficiency virus

Feline immunodeficiency virus (FIV) was first isolated in 1987 from a group of cats in Petaluma, California (Pedersen *et al.,* 1987). Infected cats within the group displayed generalised lymphadenopathy and acute neurological disease. FIV was found to exhibit many characteristics in common with HIV infection in humans and was classified as a member of the lentivirus genus (Olmsted *et al.,* 1989a, Olmsted *et al.,* 1989b, Pedersen *et al.,* 1989, Brown *et al.,* 1994, Langley *et al.,* 1994). Experimental and seroepidemiologic studies suggest that FIV is transmitted by bites and is most prevalent in high-density populations of free roaming cats (feral and pet). Kittens experimentally infected with FIV do not follow the disease course typically observed in naturally infected cats and many develop a transient fever and neutropenia 4 - 8 weeks after inoculation (Yamamoto *et al.,* 1988, Pedersen *et al.,* 1989). Under these circumstances, infection is associated with a generalised lymphadenopathy persisting for up to 9 months, after which time most cats recover. Terminal immunodeficiency is mainly found in naturally infected cats. Because of the molecular and biological similarities shared with HIV, FIV has been used as an animal model to study human immunodeficiency virus infection.

1.2.3 Simian immunodeficiency viruses

In terms of significance to immunodeficiency in humans, simian immunodeficiency viruses (SIV) represent the most important of all lentiviruses. SIVs are the closest known relatives to HIV; moreover, recent analyses have indicated that members of this group of viruses are likely to have been the progenitors of HIV (see section 1.4). Unlike HIV, within their natural hosts members of the SIV group are not associated with pathological conditions. Highly prevalent in a variety of Old World primate

species (Desrosiers & Letvin, 1987, Desrosiers *et al.*, 1990, Gojobori *et al.*, 1990, Johnson *et al.*, 1991, Myers & Pavlakis, 1992) strains of SIV are designated according to their species of origin. Naturally occurring infection has been described in sooty mangabeys (*Cercocebus atys,* SIV_{sm}), mandrills (*Papio sphinx,* SIV_{mnd}), Sykes' monkeys (*Cercopithecus mitis alboglugaris,* SIV_{syk}) and African green monkeys (*Cercopithecus* spp, SIV_{agm}), the latter group being further sub-divided into four African green monkey subspecies.

SIV has also been isolated from wild caught captive chimpanzees (SIV_{cpz}) and has led to the characterisation of four SIV_{cpz} strains that fall into two distinct subspecies groups analogous to those defining African green monkey SIV strains (Huet *et al.*, 1990, Janssens *et al.*, 1994, Vanden Haesevelde *et al.*, 1996, Gao *et al.*, 1999). Serological screening had, until recently, failed to detect significant prevalence of natural infection among wild populations of chimpanzees. However in 1999, a phylogenetic study reported the discovery of a recombinant SIV_{cpz} genome derived from divergent SIV_{cpz} lineages suggesting that screening of free-living adult chimpanzees from all four chimpanzee subspecies may reveal the existence of a natural SIV_{cpz} reservoir (Gao *et al.*, 1999) (see section 1.4).

In contrast to the described natural infections, SIV infection of another primate species that does lead to an immunodeficiency-like illness has been documented. A number of primate centres in America characterised SIV isolates from rhesus, pigtail, cynomologous, and stumptailed macaques (SIV_{mac}, SIV_{mne}, SIV_{cyn}, and SIV_{stm}, respectively) presenting with diseases such as lymphomas, progressive multifocal leukoencephalopathy, and tuberculosis. Initially the macaque isolates were considered as independent strains derived from a natural macaque SIV reservoir. Screening of feral macaques however, did not reveal any natural SIV infection and lead to the investigation into other routes of transmission. Extensive research into housing and inoculation histories in conjunction with molecular characterisation studies has indicated that cases of macaque SIV infection are examples of zoonotic transmission (Daniel *et al.*, 1985, Hirsch *et al.*, 1989a, Hirsch *et al.*, 1989b, Novembre *et al.*, 1992). When compared phylogenetically, the macaque isolates are as equidistantly related to SIV_{sm} as they are to each other. It is likely that the macaque SIV isolates arose as a result of a number of cross species transmission events that probably took place in the setting of the primate centres of the 1970's where Asian macaques are thought to have been housed with wild caught sooty mangabeys (Murphey-Corb *et al.*, 1986, Mansfield *et al.*, 1995). Ironically, the ability of the macaque SIV isolates to cause immunodeficiency-like disease in monkeys similar to AIDS in humans has provided an extremely valuable tool for the study of the pathogenesis of human lentiviruses.

1.2.4 Isolation of HIV-1

In 1981, the Centre for Disease Control (CDC) identified a new syndrome described as 'Gay Related Immune Deficiency' (GRID) in two previously healthy, young homosexual males. The two patients had presented with severe immunodeficiency characterised by profound depletion of $CD4^+$ lymphocytes, in addition to *Pneumocystis carinii* pneumonia and Kaposi's sarcoma (Control, 1981a) (Control, 1981b). Subsequently, symptoms similar to those accompanying GRID began to be described in intravenous (IV) drug users, transfusion recipients, haemophiliacs, and infants of IV drug dependent mothers. In 1982, Acquired Immunodeficiency Syndrome (AIDS) was the name chosen to describe this group of immunodeficiencies.

The epidemiology of the early AIDS epidemic was characterised by an increasing incidence and clustering of cases in large cities such as New York, suggesting the involvement of a transmissible agent. The profound depletion of $CD4^+$ T lymphocytes observed in all AIDS patients was similar to the immune deficiency identified in cases of human and animal retroviral infection. Since human retroviral infection was known to cause leukaemias, lymphomas, and T cell malignancies; several laboratories predicted that the putative causative agent for AIDS would be a retrovirus.

In 1983, credence was lent to this prediction when an American group lead by Robert Gallo reported the isolation of T-cell tropic HTLV from an AIDS patient (Gallo *et al.,* 1983). The suggestion that this retrovirus was the causative agent of AIDS was, however, inconsistent with the characteristic lytic infection of $CD4^+$ lymphocytes observed in AIDS patients; HTLV was known to immortalise T lymphocytes, not kill them. In the same issue of Science a group at the Pasteur Institute lead by Luc Montangnier and Françoise Barré-Sinoussi, reported the isolation of a novel retrovirus from a patient displaying multiple lymphadenopathies (Barré-Sinoussi *et al.,* 1983). Importantly, this retrovirus, initially termed lymphadenopathy associated virus (LAV), was found to be distinct from HTLV I, and II, and was found to infect and grow to significant titres within $CD4^+$ lymphocytes. Furthermore, LAV was found to kill rather than immortalised the $CD4^+$ lymphocytes it was cultured in.

During 1984 a further two AIDS associated retroviruses were identified, the first by Gallo and colleagues (Gallo *et al.,* 1984) and the second by an American group headed by Jay Levy (Levy *et al.,* 1984). The second AIDS associated retrovirus to be characterised by the Gallo group while distinct from HTLV, showed cross-reactivity with some HTLV-I and HTLV-II proteins, and was subsequently named HTLV-III. The retrovirus identified by Levy *et al* was recovered from healthy people as well as AIDS patients

signifying for the first time the existence of a carrier state for the AIDS virus. This retrovirus was termed AIDS-associated retrovirus (ARV).

Shortly after their identification, all three AIDS-associated retroviruses (LAV, HTLV-III and ARV) were found to exhibit similar *in vitro* cytopathic characteristics and some antigenic cross-reactivity to each other. This lead to them being recognised as members of the same group of retroviruses, termed the lentiviridae. Distinct from HTLV, the group of three retroviruses were named Human Immunodeficiency Virus (HIV) in 1986 by the International Committee on Taxonomy of Viruses (Coffin *et al* 1986). In later studies the HTLV-III strain, isolated by the Gallo group was confirmed to be the same virus as LAV, while ARV is considered a separate strain of HIV-1. After extensive isolation of HIV from the blood of many patients with AIDS and ARC, as well as from the PBMC of several clinically healthy individuals, the causal link between HIV and AIDS is now widely accepted.

1.2.5 Identification of HIV-2

Following the identification of HIV-1 and its associated pathogenesis the alarming consequence of a worldwide pandemic was the subject of intense debate and investigation. Large serological screening studies were initiated in many countries to determine the prevalence of HIV infection. Cases of HIV-1 had been widely reported within Central and Eastern Africa however few cases had been identified in West African countries. In 1985, Barin *et al* published data produced from a study investigating HIV prevalence in healthy female prostitutes attending an STD clinic in Senegal, West Africa (Barin *et al.,* 1985). Sera taken from the Senegalese individuals were examined for reactivity to HIV-1 and SIV antigens. Surprisingly the group produced stronger serological responses against SIV_{mac} antigens than against HIV-1 viral antigens, indicating that the virus that the Senegalese prostitutes had been exposed to was related to HIV-1, but exhibited a closer relationship to SIV. Following this discovery, in 1986 Clavel and co-workers isolated a new human retrovirus from the peripheral blood mononuclear cells (PBMC) of two West African patients exhibiting an AIDS-like illness (Clavel *et al.,* 1986a). Analysis of this retrovirus showed it to be more serologically and antigenically related to SIV than HIV-1, it was therefore termed human immunodeficiency virus type 2 (HIV-2) and became the prototypic HIV-2 strain ROD. Molecular cloning and sequencing of the newly isolated retrovirus revealed the explanation for the stronger serological response to SIV antigens reported in both the Clavel and Barin studies (Clavel *et al.,* 1986b). The HIV-2 genome isolated by Clavel and co-workers displayed only 42% homology with the HIV-1 genome, but showed 75% similarity to strains of SIV_{sm} and SIV_{mac}. Further molecular and

serological examination of HIV-2 has since confirmed that it is more closely related to SIV strains than strains of HIV-1 (Chakrabarti *et al.*, 1987, Franchini *et al.*, 1987, Hirsch *et al.*, 1989, Gao *et al.*, 1992, Gao *et al.*, 1994, Chen *et al.*, 1996, Chen *et al.*, 1997) (see section 1.4).

Since early studies of HIV-2 infection demonstrated little association with disease there was doubt as to whether HIV-2 was an etiologic agent of AIDS (Barin *et al.*, 1985, Kanki *et al.*, 1987). However, in 1987 a study investigating the clinical status of 30 HIV-2 seropositive individuals revealed that HIV-2 was the causative agent of AIDS in 17 of the patients, while 6 remained asymptomatically infected. Further serological screening studies have since identified HIV-2 as the dominant HIV-type circulating West Africa.

1.3 Primate lentivirus phylogeny

The genetic relationships between members of the primate lentivirus family have been established by phylogenetic analysis of the three major genes within the retrovirus genome, *gag, pol,* and *env.* Differences within these genes occur as a result of the lack of proofreading mechanisms in the retroviral reverse transcriptase and give rise to the diversity seen within the HIV-1, HIV-2, and SIV lineages. Mutation rates as high as 1×10^4 mutations per genome per replication cycle determine the pleomorphic nature of the primate lentiviral strains and can confer selective replication advantages and disadvantages (Preston *et al.,* 1988, Roberts *et al.,* 1988). Based on comparative genetic sequence analysis five distinct groups of primate lentiviruses have been identified. These are HIV-1/SIV_{cpz}, HIV-2/SIV_{sm}/SIV_{mac}, SIV_{agm}, SIV_{mnd}, and SIV_{syk} (Figure 1.1A, pg 38).

Three groups contain strains isolated from three Old World African primate species, SIV_{agm}, SIV_{mnd}, and SIV_{syk}. The remaining two groups contain a mixture of strains taken from human and primate species and reflect the finding that HIV-1strains are more related phylogenetically to SIV strains derived from the chimpanzee, while HIV-2 strains phylogenetically cluster with strains of SIV_{sm}.

Phylogenetic analysis of specific gene sequences further subdivides viral strains belonging to each of these groups into subtypes or clades. While approximately 40% divergence is observed between the five primate lentiviral groups; viral subtypes within these groups are approximately 10-20% divergent from each other.

In addition to RT-induced strain diversity, genetic diversity is introduced into each of the five primate lentiviral lineages by a process termed subtype recombination. Strains, which have been generated by the genetic recombination of two distinct viral genomes, have been identified in each of the HIV-1, HIV-2 and SIV lentiviral families. These mosaic viral strains indicate that superinfection is a process that does occur within each of these lentiviral infections.

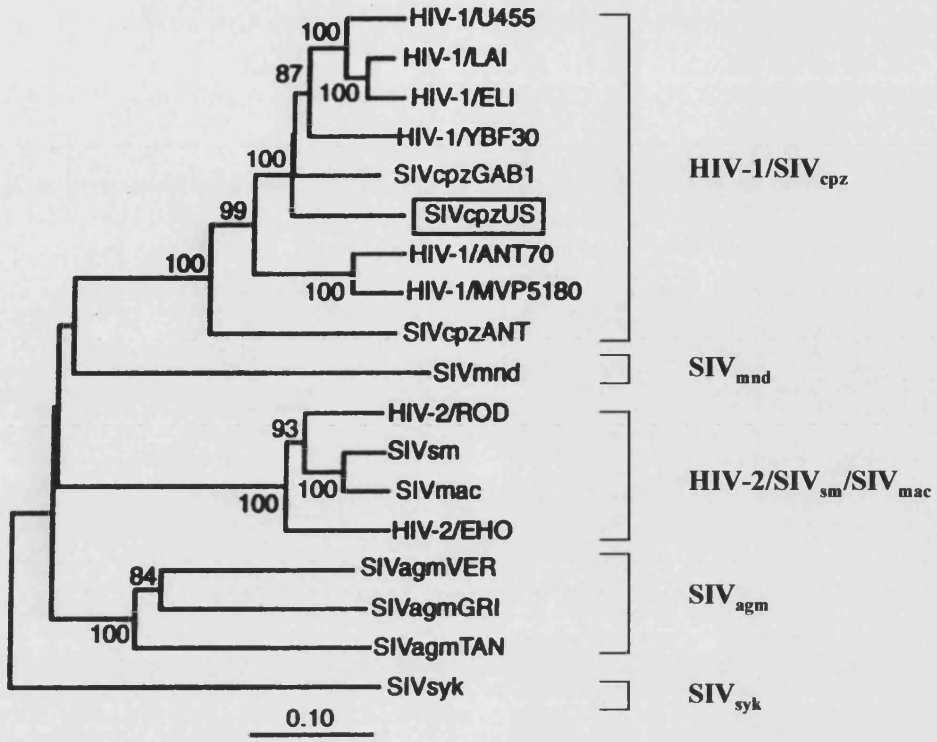
Phylogenetic analysis of *gag* gene sequences has lead to the division of HIV-1 strains into three distinct phylogenetic groups. The first major group to be classified contains the majority of HIV-1 isolates and is termed the main group or group M. Isolates within this group have been subdivided further into ten distinct but phylogenetically related subtypes or clades, A-J (Louwagie *et al.,* 1993, Myers *et al.,* 1993). Circulating HIV-1 strains identified in North America and Europe largely fall into subtype B, whereas

subtype A and D viruses have been predominantly isolated from Central Africa. Isolates identified in Thailand have primarily been characterised as subtype E, and subtypes F to J represent small clusters of isolates from patients with links to Central African countries, in particular Zaire (Kostrikis *et al.,* 1995, Leitner *et al.,* 1995). Strains within subtypes B and D share the highest degree of homology indicating that HIV-1 circulating America and Europe may have originated from Central Africa.

The identification of two highly divergent HIV-1 isolates derived from patients originating from the Cameroon lead to the classification of a second HIV-1 phylogenetic group, the "outlier" group – O (De Leys *et al.,* 1990, Gurtler *et al.,* 1994, Vanden Haesevelde *et al.,* 1994). This group now contains a number of other isolates exhibiting significant divergence from members of the group M subtypes; the vast majority also originate from Cameroon (Charneau *et al.,* 1994, Loussert-Ajaka *et al.,* 1995). A third phylogenetic group has most recently been defined and represents strains of HIV-1 which fail to fall into the M or O group classification. The group termed N (non-M/non-O) is the least widespread of all HIV-1 lineages and thus far contains two highly divergent isolates derived from two Cameroonian patients.

Just as strains of HIV-1 exhibit high levels of divergence, isolates of HIV-2 have been shown to display an equivalent range of divergent genomes (Dietrich *et al.,* 1989, Kuhnel *et al.,* 1989, Tristem *et al.,* 1989). Phylogenetic analysis of HIV-2 *gag, pol, env,* and LTR sequences taken from isolated HIV-2 genomes has identified six distinct HIV-2 subtypes, A-F (Figure 1.1B, pg 38). The majority of HIV-2 isolates identified so far cluster within what has been termed subtype A. The second largest group of HIV-2 isolates form subtype B, isolates within this group display greater than 20% divergence to strains within subtype A. The first divergent members of this group to be identified were isolates, D205 and GH-1, isolated from two Ghanaian patients, one asymptomatic the other displaying AIDS related symptoms (Dietrich *et al.,* 1989, Miura *et al.,* 1990). Following the identification of the divergent isolates in subtype B, Gao and co-workers analysed the infecting HIV-2 strains within 12 patients originating from six different West African countries (Gao *et al.,* 1994). Phylogenetic analysis of the isolates identified in this study defined a further three HIV-2 subtypes, C, D, and E. Detailed genetic analysis of one particular strain HIV-2_{F074} showed that *pol, env,* and LTR regions PCR-amplified from this patient were virtually identical to the corresponding regions within isolates of SIV_{sm}. This strain, isolated from a healthy Liberian worker, was placed in subtype D, and as described in section 1.4, provided evidence for an SIV progenitor of HIV-2. A sixth subtype of HIV-2, termed F, was recently identified following the isolation of a highly divergent strain, 93SL2 (Chen *et al.,* 1997a). This strain was found as a single occurrence in Sierra Leone, a region in which household pet sooty mangabeys are common.

A



B

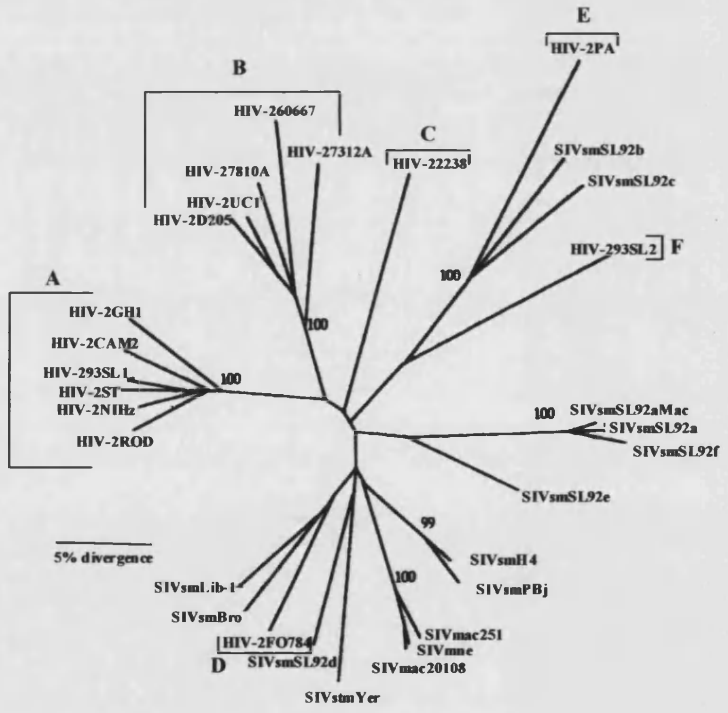


Figure 1.1 Neighbour-joining trees showing primate lentiviral phylogenetic relationships. (A). Primate lentiviral Pol sequence analysis. Brackets to the right indicate grouping of the five primate lentiviral lineages (adapted from Gao *et al.*, 1999). (B). HIV-2 and SIV Gag sequence analysis showing the six subtypes, A-F, and clustering of HIV-2 F0784 with SIV (adapted from Chen *et al.*, 1997).

1.4 Origins of HIV

As more isolates of HIV-2 were obtained, mounting evidence showing the remarkable similarities to SIV raised the possibility that HIV-2 infection in humans had arisen as a result of a zoonotic transmission of SIV from monkey to man. A number of studies utilising nucleotide sequence analyses and RNA hybridisation technology investigated the potential primate candidates, primarily centring upon SIV_{agm}, SIV_{mac}, and SIV_{sm}. As outlined in Figure 1.1A, pg 38), phylogenetic analysis revealed that strains of HIV-2 clustered with SIV_{sm} and SIV_{mac}, while SIV_{agm} strains formed a less related distinct cluster (Hirsch *et al.*, 1989). The clustering of SIV_{mac} strains with strains of SIV_{sm} was established as resulting from a cross-species transmission of SIV_{sm} from infected sooty mangabeys to Asian macaques in the setting of the Primate Research Centres in the 1970s (Murphey-Corb *et al.*, 1986, Mansfield *et al.*, 1995), leaving SIV_{sm} as the only known natural simian candidate for monkey to man zoonosis. Restriction mapping of SIV_{sm} and HIV-2 genomes went on to demonstrate that in some cases HIV-2 isolates were as closely related to SIV_{sm} isolates as to other HIV-2 isolates (Gao *et al.*, 1992, Chen *et al.*, 1996).

It is thought that between 20 - 30% (Chen *et al.*, 1996) of the indigenous free-living West African sooty mangabey population is infected with SIV_{sm}. This population coincides with the geographical distribution of documented HIV-2 infection in West Africa and since the sooty mangabey monkey is hunted for food and kept as a domestic pet throughout West Africa the opportunity for cross-transmission is evident. A defining piece of evidence implicating the zoonotic transmission of SIV_{sm} to humans was discovered following the detailed analysis of the HIV-2 subtype D strain $HIV\text{-}2_{F0784}$, identified by Gao and colleagues in 1994 (Gao *et al.*, 1992, Gao *et al.*, 1994). Several regions of the F0784 genome were found to be virtually identical to corresponding regions within stains of SIV_{sm}. Moreover, with the addition of strains like $HIV\text{-}2_{F0784}$ to phylogenetic analysis, in 1994 Sharp *et al* demonstrated that HIV-2 and SIV_{sm} strains could no longer be separated into distinct phylogenetic lineages and should be considered as a single genetically diverse group of viruses (Sharp *et al.*, 1994). In order to account for the existence of the genetic diversity exhibited by the strains within the HIV-2 subtypes it is thought that up to six independent cross-species transmissions of genetically diverse SIV strains to human populations have occurred. In accordance with this, a seroprevalence study carried out by Chen and co-workers (Chen *et al.*, 1997) has proved that substantial genetic divergence exists within the feral and household mangabey population that could account for the cross-species divergence observed between the subtypes

of HIV-2. To conclude, in 1994 the accidental transmission of SIVsm to a laboratory worker showed that humans are susceptible to infection by SIV_{sm}.

Up until 1999 evidence for an SIV progenitor of HIV-1 had been less conclusive, although isolates from chimpanzees (SIV_{cpz}) had been shown to be antigenically closer to HIV-1 than they were to HIV-2 or other SIV strains (Petters *et al.*, 1989, Huet *et al.*, 1990, Janssens *et al.*, 1994, Vanden Haesevelde *et al.*, 1996). Criticisms levelled at the chimpanzee as a natural host and reservoir for HIV-1 centred around the wide spectrum of diversity between HIV-1 and SIV_{cpz}, and the apparent low prevalence of SIV_{cpz} infection in wild living animals. Up until this point only three other naturally occurring SIV_{cpz} infections had been identified ($SIV_{cpz\ GAB1}$, $SIV_{cpz\ GAB2}$, $SIV_{cpz\ ANT}$) (Huet *et al.*, 1990, Janssens *et al.*, 1994, Vanden Haesevelde *et al.*, 1996). In addition, chimps were found to be present in geographic regions of Africa (Teleki, 1989) where AIDS was not initially recognised.

A recent study by Gao *et al,* however, has provided the most persuasive evidence yet that HIV-1 was transmitted to humans from the chimpanzee, *Pan troglodytes* harbouring SIV_{cpz} (Gao *et al,* 1999). In 1999, the group at the University of Alabama identified a fourth chimp exhibiting natural SIV_{cpz} infection. PCR amplification and sequence analysis revealed that the virus isolated from the fourth chimp was related to but distinct from known SIV_{cpz} and HIV-1 strains, and was named $SIV_{cpz\ US}$. Phylogenetic analysis demonstrated that the four naturally occurring SIV_{cpz} strains fell into two related yet highly divergent subspecies-specific phylogenetic lineages. The main group contained $SIV_{cpz\ GAB1}$, $SIV_{cpz\ GAB2}$, and $SIV_{cpz\ US}$, all derived from the *Pan troglodytes troglodytes* subspecies of chimpanzee, while $SIV_{cpz\ ANT}$ fell into an outlier group and was obtained from a member of the *Pan troglodyte schweinfurthil* subspecies. This finding indicated that, as observed in SIV_{agm}, SIV_{cpz} had undergone host dependent evolution. Following this discovery, in an attempt to identify any cross-transmission events, Gao and co-workers re-examined the phylogenetic positions of all known strains within the SIV_{cpz} and M, N, and O HIV-1 groups. All HIV-1 strains were found to phylogenetically cluster with SIV_{cpz} strains infecting *Pan troglodytes troglodytes,* a primate whose natural host range coincides precisely with areas of HIV-1 group M, N, and O endemicity. This finding, in conjunction with the detection of subspecies-dependent evolution and recombination between SIV_{cpz} lineages, provides substantial evidence that wild living chimps must have been a widely infected long-standing natural reservoir of SIV_{cpz}, but also indicates that HIV-1 groups M, N, and O have arisen as a consequence of independent zoonotic transmissions of SIV_{cpz} from *Pan troglodytes troglodytes* to man (Gao *et al,* 1999, Hahn *et al.*, 2000). Initial screenings of free-living adult chimps have indicated that the prevalence of SIV in chimps is far lower (>2%) than in other

primates such as African green monkeys, where it can reach 90% in adults (Santiago *et al.,* 2002). However, a recent report by Santiago *et al* have identified a Tanzanian chimp population showing SIV prevalences of between 10 – 15% (Santiago *et al.,* 2003). Only extensive screening of free-living adult chimps from all four subspecies, as well as human populations from corresponding geographic localities, will determine the extent of natural SIV_{cpz} infection and give an indication of the frequency of zoonotic transmission to humans. Such studies would be able to truly verify that SIV_{cpz} is the natural host and reservoir for HIV-1.

1.5 Natural history of HIV-2 infection

1.5.1 Epidemiological aspects of HIV-2 infection.

Following the initial isolation of HIV-2 from West African individuals (Clavel *et al.*, 1986a, Clavel *et al.*, 1987a) large-scale immunological studies have revealed that Human Immunodeficiency Virus type 2 (HIV-2) is endemic in, although largely restricted to, the many countries of West Africa (Decock & Brun Vezinet, 1989). Community-based studies have shown that of all the West African countries, Guinea-Bissau has the highest serological prevalence of HIV-2 (Poulsen *et al.*, 1989, Naucler *et al.*, 1999, Wilkins *et al.*, 1993, Ricard *et al.*, 1994). Women attending mother-and-child health clinics in the Bafata province of this country reported an HIV-2 prevalence of 9.2%. Surrounding countries including The Gambia, Senegal, The Ivory Coast, and Ghana have a lower HIV-2 seroprevalence ranging from 1-2% (Obisesan *et al.*, 1997, Kassim *et al.*, 1998, Whittle *et al.*, 1998, UNAIDS & WHO, 2000). Nevertheless, strikingly higher prevalences have been observed within the commercial sex worker (CSW) populations of these countries. CSW HIV-2 prevalences range from 8.1% in Mali (Peeters *et al.*, 1998) to an alarming 27.5% in The Gambia (Hawkes *et al.*, 1995, Langley *et al.*, 1996, Ghys *et al.*, 1997, Samb *et al.*, 1997).

A number of cases of HIV-2 infection have also been described in countries with socio-economic links to West Africa including the United Kingdom, France, Portugal, and the south-western regions of India and South America (Rey *et al.*, 1987, Cortes *et al.*, 1989, Smallman-Raynor & Cliff, 1991, Rubsamen Waigmann *et al.*, 1991, Grez *et al.*, 1994, Quinn *et al.*, 1994). In addition, isolated cases have been described in several European countries (Doerr *et al.*, 1987, Evans *et al.*, 1991, van der Ende *et al.*, 1996) and the USA (Ayanian *et al.*, 1989, O'Brien *et al.*, 1992). The vast majority of cases identified outside of West Africa however, are either Europeans with West African sexual partners or are West African themselves (Smallman-Raynor & Cliff, 1991).

1.5.1.1 Age distribution

In contrast to STDs such as HIV-1 whereby the age peak distribution of infections lies within the young adult age group of 20-34 years (Killewo *et al.*, 1990, Wawer *et al.*, 1991, Wagner *et al.*, 1993), the vast majority of individuals infected with HIV-2 fall into age categories of 45 years and above. A number of

studies have outlined that HIV-2 prevalence in women is highest in age groups 35-45 years (Wilkins *et al.*, 1991b, Wilkins *et al.*, 1993), or 50-59 years (Poulsen *et al.*, 1989), and in men 50-59 years (Poulsen *et al.*, 1989, Wilkins *et al.*, 1991b, Wilkins *et al.*, 1993). The most likely explanation for the difference in age distribution between the two infections lies in the fact that people infected with HIV-2 have a greater chance of survival and live to an older age than individuals infected with HIV-1. This issue will be further discussed in section iv).

1.5.1.2 Transmission of HIV-2

The increased incidence and clustering of early AIDS cases within the homosexual, heterosexual, intravenous drug user and blood transfusion recipient populations gave the first indications that an infectious transmissible agent was involved in the AIDS epidemic. Transmission routes, and the risk factors involved are similar for both HIV-1 and HIV-2 infection. Both viruses can be transmitted by homosexual (Gottlieb & Ackerman, 1982, Jaffe *et al.*, 1983, Brucker *et al.*, 1987, Pieniazek *et al.*, 1991, Soriano *et al.*, 1996) and heterosexual contact (Harris *et al.*, 1983, Gao *et al.*, 1993, Aaby *et al.*, 1996, Ghys *et al.*, 1997), perinatal transmission (Adjorlolo-Johnson *et al.*, 1994, Cavaco-Silva *et al.*, 1998), transfusion of blood and blood products (Dufoort *et al.*, 1988, Poulsen *et al.*, 1989, Faye *et al.*, 1997), and intravenous needle sharing. However, few studies exist demonstrating the latter two modes of transmission for HIV-2. In West Africa, the principal route by which HIV-2 is transmitted is via heterosexual contact. The migration of female prostitutes across the region, both to and from areas of high endemicity has played a significant role in the spread of HIV-2 within the cities and rural communities of the countries in West Africa (Wilkins *et al.*, 1991a, Pickering *et al.*, 1992). As have events such as the colonial occupation of Guinea Bissau by the Portuguese, which largely account for the higher prevalence of HIV-2 within this country. However, a number of studies have reported that the rate of male to female transmission of HIV-2 is between 5- and 9-fold lower than that observed with HIV-1 (Schim van der Loeff & Aaby, 1999). The reduced heterosexual transmission rate observed in HIV-2 infection is thought to be related the very low plasma viral loads observed within the majority of asymptomatic HIV-2-infected individuals (Adjorlolo-Johnson *et al.*, 1994, Kanki *et al.*, 1994, Berry *et al.*, 1998, Cavaco-Silva *et al.*, 1998, O'Donovan *et al.*, 2000) (see section v). An investigation into the level of viral RNA (vRNA) in the cervico-vaginal secretions (CVS) of female CSWs in Cote d'Ivoire found virus in only 5% of HIV-2-infected women as opposed to 24% of women infected with HIV-1 (Ghys *et al.*, 1997). Similar findings were also identified in female CSWs from Dakar, Senegal (Samb *et al.*, 1997, Sankale *et al.*, 1998).

In addition to the rate of heterosexual transmission, vertical transmission of HIV-2 occurs at a much lower rate than documented for HIV-1 (Adjorlolo-Johnson *et al.*, 1994, O'Donovan *et al.*, 2000). Vertical transmission rates for HIV-1 in the absence of antiretroviral therapy range from between 15-35%. In contrast, the estimated rate of vertical HIV-2 transmission is between 0-4% (Del Mistro *et al.*, 1992, Gayle *et al.*, 1992, Andreasson *et al.*, 1993, Cavaco-Silva *et al.*, 1998, O'Donovan *et al.*, 2000). A recent study by O'Donovan and co-workers demonstrated that the lower plasma viral loads of HIV-2 seropositive mothers when compared to HIV-1-infected mothers related to the significantly reduced rate of mother to child transmission observed in HIV-2 infection (O'Donovan *et al.*, 2000). The reduced risk of transmission that is associated with HIV-2 is thought to account for the slower spread of HIV-2 infection in comparison to HIV-1. This has given rise to an endemically infected population largely confined to West Africa in contrast to the rapidly spreading HIV-1 pandemic.

1.5.1.3 Clinical aspects of HIV infection

Within one to four weeks of primary infection with HIV-1 or HIV-2 a significant proportion of individuals experience an acute clinical illness, which manifests in influenza-like symptoms accompanied by high levels of virus detectable in the blood, 10^5 to 10^7 virions/ml (Gaines *et al.*, 1987, Clark *et al.*, 1991, Daar *et al.*, 1991). Documented primary infection symptoms consist of headache, muscle aches, sore throat, fever, swollen lymph nodes and rash (Cooper *et al.*, 1985, Tindall *et al.*, 1988, Besnier *et al.*, 1990). The symptoms of acute infection normally persist for two to three weeks but usually result in clinical recovery. During the acute period of infection high levels of proinflammatory cytokines associated with immune activation are expressed. The increase in the levels of cytokines such as interleukin-1β (IL-1β), tumour necrosis factor alpha (TNF-α) and gamma interferon (IFN-γ), is thought to be associated with the significant $CD8^+$ cell expansion that is observed in the second week of primary infection (Cooper *et al.*, 1988, Pantaleo *et al.*, 1994a, Pantaleo *et al.*, 1994b). The appearance of this cytotoxic T lymphocyte (CTL) response is accompanied by a significant reduction in the level of infectious virus within the peripheral circulation, however, the detection of low-level viremia following this stage of infection indicates that HIV variants do escape this antiviral cellular immune response. Several studies have demonstrated that the efficiency of viral clearance and the subsequent quantity of peripheral viral HIV RNA remaining after primary infection, termed the "set point", are predictive of the length and severity of the disease course in both HIV-1 and HIV-2 infections (Keet *et al.*, 1993, Mellors *et al.*, 1995, Wei *et al.*, 1995, Andersson *et al.*, 2000). Following the cell-mediated immune response, a broad spectrum of antiviral antibodies can be detected within the infected individual, signifying

seroconversion. Neutralising antibodies targeting HIV have been detected within 20 days after virus transmission, but in general appear approximately ten weeks post-infection.

The acute phase of HIV infection is consequently followed by a clinically asymptomatic period, which can persist for several years. It is this phase of infection that largely distinguishes HIV-1 from HIV-2 infection (see section iv). During the asymptomatic period, molecular analysis has revealed that transcriptionally active virus is present, directing the low level expression of virion RNA, which, in the case of HIV-1, can be detected at all stages of infection (Michael *et al.,* 1992, Piatek *et al.,* 1993, Saksela *et al.,* 1995). The asymptomatic phase of HIV infection is characterised by a very gradual lymphadenopathy. Early studies of HIV infection revealed that while a wide variety of human cells were susceptible to HIV, the $CD4^+$ T-helper lymphocyte was the major target of HIV infection; $CD4^+$ cells replicated HIV to the highest titres. It is the linear decline of this subset of immune cells over time, which eventually results in the clinical manifestation of immunodeficiency. A variety of symptoms herald the deterioration of the immune system and the onset of AIDS or AIDS-related complex (ARC) whether caused by infection with HIV-1 or HIV-2 (Mabey *et al.,* 1988, Naucler *et al.,* 1989, Decock *at al.,* 1990, Grant *et al.,* 1997, Poulsen *et al.,* 1997). These include chronic fevers, diarrhoea, weight loss, oral thrush, and hairy leukoplakia. The presence of two or more of these symptoms characterise ARC and leads to the development of AIDS. Once the diagnosis of AIDS has been made, survival of the patient is often less than two years unless the patient is treated with antiretroviral drugs.

Patients in the symptomatic phase of infection often present with a number of opportunistic infections caused by pathogens classically associated with severe immunodeficiency (Mildvan *et al.,* 1982). These are broadly similar for both HIV-1 and HIV-2 infections and include *Pneumonocystis carinii* pneumonia (PCP), *M. tuberculosis*, CMV, and malignancies such as Kaposi's sarcoma, B-cell lymphoma and rectal cancers. A recent study by Ariyoshi *et al* found however, that Gambian patients infected with HIV-2 were 12.4 times less likely to have Kaposi's sarcoma than patients infected with HIV-1 (Ariyoshi *et al.,* 1998).

Several opportunistic infections can present with neurological manifestations, and a significant proportion of AIDS patients display an HIV-associated subacute encephalitis termed AIDS dementia complex. The mean survival time from the onset of severe dementia is usually less than six months.

1.5.1.4 Pathogenicity of HIV-2

Although a causal link between HIV-2 and AIDS was proved soon after the initial characterisation (Clavel *et al.*, 1986, Brun-Vezinet *et al.*, 1987, Albert *et al.*, 1989), relatively few cases of HIV-2 associated disease were identified in preliminary studies (Barin *et al.*, 1985, Kanki *et al.*, 1987). Following on from this, early community based studies demonstrated that mortality rates among HIV-2 infected individuals were only two- to three-fold higher than rates among uninfected controls (Ricard *et al.*, 1994, Poulsen *et al.*, 1997). Similar studies performed with HIV-1 infected individuals had shown a rate of mortality ten-fold that of uninfected controls (Morgan *et al.*, 1997, Nunn *et al.*, 1997, Todd *et al.*, 1997). Therefore, in the vast majority of cases, infection with HIV-2 was thought to lead to a much better disease prognosis than infection with HIV-1.

Since then a number of comparative epidemiological studies have confirmed that progression to immunosuppression and AIDS following HIV-2 infection is slower than for HIV-1, and is associated with a relatively prolonged clinically asymptomatic period (Pepin *et al.*, 1991, Marlink *et al.*, 1994, Ricard *et al.*, 1994, Lisse *et al.*, 1996). In fact, due to the prolonged maintenance (>56 months) of low to undetectable plasma vRNA (<500 copies/ml), high CD4 cell percentages of >28%, and the lack of symptoms of clinical progression, the vast majority of individuals infected with HIV-2 are categorised as long-term-non-progressors (LTNP). In a comparison of the rate of CD4 decline over a four year period, Jaffar *et al* found that 47% of HIV-2 infected patients, versus 32% of HIV-1 infected individuals, showed no decline in CD4 cell percentage (Jaffar *et al.*, 1997). In addition, in cases where CD4 cell percentage decline was seen, the rate of decline was significantly slower in HIV-2 patients. A recent community-based study performed in the remote village of Caio, Guinea-Bissau, indicated that the mortality rate of HIV-2-infected subjects was only 1.7-fold greater than that of HIV-uninfected controls (Berry *et al.*, 2002). Moreover, cases of non-progressive asymptomatic periods extending in excess of 27 years have been reported; HIV-2 infection in individuals over and above the age of 60 years is not uncommon (Ariyoshi *et al.*, 1996).

The length of the symptomatic phase of HIV-2 infection is also thought to be longer than in HIV-1 infection. HIV-2 AIDS patients in a Gambian hospital-based study were found to survive up to three times longer than HIV-1 AIDS patients (Whittle *et al.*, 1994); similar observations were also made in a study of African HIV-2 AIDS patients living in Paris (Matheron *et al.*, 1997). The lower rate of transmissibility and the reduced risk of mortality observed in HIV-2 infection have lead to the belief that

HIV-2 is a less pathogenic virus than HIV-1 (Marlink *et al.*, 1994, Whittle *et al.*, 1994, Jaffar *et al.*, 1997).

Nevertheless, a number of HIV-2-infected patients displaying relatively rapid progression to AIDS, both in The Gambia and among West African residents in Europe have been identified (van der Ende *et al.*, 1996, Ariyoshi *et al.*, 1998). The relatively small number of cases termed rapid progressors (RP) display a much more "HIV-1-like" disease course, and tend to have been identified in younger individuals i.e. those under the age of 45 years (Poulsen *et al.*, 1989, Ricard *et al.*, 1994, Norrgren *et al.*, 1995, Poulsen *et al.*, 1997). The rate of mortality among HIV-2-infected adults under the age of 45 years has been shown to be 5:1 (Norrgren *et al.*, 1995, Poulsen *et al.*, 1997), significantly higher than the mortality ratio of older HIV-2 infected individuals, suggesting a dichotomy between the progression rates in HIV-2 infection (Jaffar *et al.*, 1997, Grassly *et al.*, 1998).

The length of time between acute primary infection and the onset of symptomatic HIV infection within an individual is influenced by a variety of viral and host factors. However, data comparing the levels of peripheral blood viral HIV-2 and HIV-1 RNA has provided evidence suggesting that the slower rate of progression to symptomatic disease observed in the majority of HIV-2 infections may result from a reduced level of viral production during the asymptomatic phase of HIV-2 infection compared to HIV-1 (Simon *et al.*, 1993, Berry *et al.*, 1998, Popper *et al.*, 2000, Berry *et al.*, 2002). In 1998, Berry and co-workers demonstrated that HIV-1 and HIV-2 infection, could be differentiated by RNA plasma viral load at the early stage of disease (Berry *et al.*, 1998). RNA virus was detectable in the plasma of all HIV-1 infected individuals with a CD4% of >28, but was detected in only one third of matched HIV-2 infected subjects. In addition, a study by Andersson *et al* demonstrated that the plasma virus set point, i.e. the semi-equilibrium reached following seroconversion, was 28-fold lower in HIV-2 seroconverters compared to HIV-1 seroconverters (Andersson *et al.*, 2000). It has been demonstrated that base-line HIV-1 and HIV-2 RNA load significantly predicts the rate of disease progression as determined by CD4+ decline or by death (Ariyoshi *et al.*, 2000, Berry *et al.*, 2002, Alabi *et al.*, 2003). Since the level of proviral DNA has been found to be similar in HIV-1 and HIV-2 infections (Berry *et al.*, 1994, Ariyoshi *et al.*, 1996, Norrgren *et al.*, 1997b, Berry *et al.*, 1998, Sarr *et al.*, 1999, Popper *et al.*, 2000), it has been suggested that the lower plasma viral loads observed in the asymptomatic phase of HIV-2 infection reflect a lower rate of virus production in comparison to HIV-1 that results in the reduced pathogenicity of HIV-2 (Popper *et al.*, 1999, Popper *et al.*, 2000). To date however, the virological and/or immunological factors involved in determining the difference between the rates of virus production in HIV-1 and HIV-2

infected individuals, and between HIV-2 infected individuals displaying slow progression and relatively rapid progression to disease have yet to be determined.

1.5.1.5 Treatment of HIV-2 infection

Due to the expense of antiretroviral treatment large-scale clinical trials involving the HIV-2 infected cohorts of West Africa have not been performed. However, a number of HIV-2 infected individuals in Europe have been given anti-retroviral agents (Soriano *et al.*, 2000, Van der Ende *et al.*, 2000, Smith *et al.*, 2001). In these cases, the principles of highly active antiretroviral therapy (HARRT) seem to apply, whereby drug combination therapy provides an effective reduction in viral load, while single or suboptimal combination therapies can lead to drug resistance (Rodes *et al.*, 2000, Van der Ende *et al.*, 2000, Smith *et al.*, 2001). It is thought that the reduced virus load and virulence associated with HIV-2 infection (see section v.) may make HARRT extremely effective within the HIV-2 infected population (Whittle *et al.*, 1998b, Smith *et al.*, 2001). Nevertheless, since the vast majority of infections occur within developing countries with very low health budgets it is unlikely that these HIV-2 infected individuals will receive any treatment without aid intervention from the countries within the developed world. The distribution of antiretroviral therapy, education, treatment and prevention of STDs, and the development of an inexpensive yet effective vaccine for the HIV-2 infected populations of the developing world remain paramount.

1.6 Genomic organisation of HIV-1 and HIV-2.

In contrast to their different pathogenicities, the genomic organisation of HIV-1 and HIV-2 is remarkably similar, as is the function of the gene products encoded by each of the genomes (Guyader *et al.,* 1987). Nucleotide sequence analysis has revealed that the HIV-2 viral genome is approximately 9.8kb in size and consists of the typical retroviral LTR-*gag-pol-env*-LTR organisation (Figure 1.2, pg 50). As well as the three contiguous structural genes and the long terminal repeat sequences that direct and regulate the expression of the viral genome, the HIV-2 and SIV genomes encode several other regulatory and accessory proteins; *tat* (*trans*activator of transcription), *rev* (regulator of expression of virion proteins), *nef* (negative factor), and viral proteins *vpr, vif,* and *vpx.* The HIV-1 genome fails to encode *vpx*, and instead contains viral protein *vpu*, which is not present in strains of HIV-2 and SIV. The primary transcript of HIV is a full length viral mRNA, which is translated into a 160kDa Gag-Pol precursor. As these two genes do not occupy the same reading frame a ribosomal frame shift is required during translation to prevent recognition of the *gag* termination codon (Jacks *et al.,* 1988). The read-through process, which generates the single Gag-Pol protein, is about 5% efficient and results in the Gag protein outnumbering the Gag-Pol protein by approximately 20-fold (Oroszlan & Luftig, 1990). During virion maturation proteolytic cleavage of the Gag-Pol precursor generates viral enzymes; p64, p53 (RT) reverse transcriptase, p34 (IN) integrase, and p11 (PR) protease. The Gag gene also encodes the Gag precursor protein (p55), proteolytic cleavage of which generates the smaller structural proteins of the virion: p24 (CA) Capsid, p17 (MA) Matrix, p9 and p7 (NC) Nucleocapsid. The Env polyprotein is expressed by another means: splicing generates a shorter subgenomic messenger RNA, gp160. This singly spliced precursor is cleaved by protease, and gives rise to the surface (SU) and transmembrane (TM) glycoproteins. Precursor mRNAs for the remaining regulatory and accessory proteins encoded by the HIV genome are produced by additional multiple splicing events.

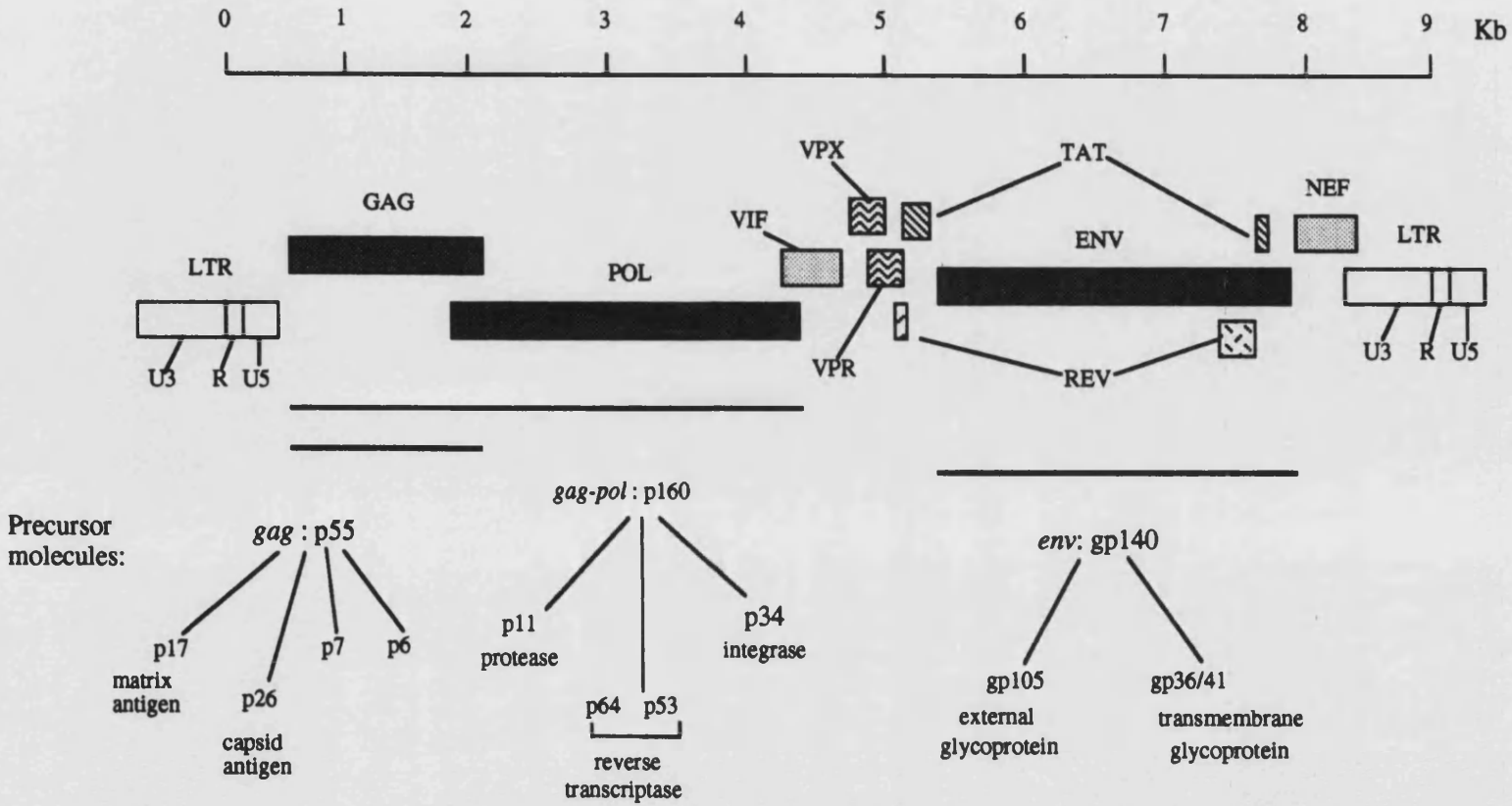


Figure 1.2 Genomic organisation of HIV-2. SIV shows similar organisation, however, HIV-1 encodes viral protein Vpu and not Vpx which is exclusive to the HIV-2 and SIV genomes.

1.7 Virion Morphology

Initially, high-resolution electron microscopy indicated that HIV virions were roughly spherical in structure; however, more recent scanning electron micrographs and computer simulation have indicated that virions are icosahedral particles about 110nm in diameter (Marx *et al.,* 1988, Nermut *et al.,* 1993). Mature virions consist of a host derived lipid bilayer envelope encapsulating the myristolated matrix protein p17 (MA), this protein in turn surrounds the characteristic type C lentiviral cone-shaped nucleocapsid comprised of viral p24 (CA), a polypeptide approximately 240 amino acids (aa) in length (Figure 1.3, pg 52). The matrix protein is associated with the lipid membrane via a signal sequence encoded within the first 31aa of the 130aa polypeptide. This region of MA also encodes a nuclear localisation signal thought to be important during the early stages of replication (Martin & Freed, 1996). Myristolation of the MA protein is required for the production of infectious virus (Bryant & Ratner, 1990).

The nucleocapsid core of the HIV virion contains two identical single-stranded molecules of genomic RNA that are closely associated with nucleocapsid protein p7 (NC-p7). The interaction between the RNA strands and p7 is determined by two cysteine-histidine rich zinc-finger like domains contained within this hydrophilic protein (Gorelick *et al.*, 1988, 1990). Also packaged into the nucleocapsid core are the *pol*-encoded products integrase (IN-p32), protease (PR-p10) and the viral RNA-dependent DNA polymerase or reverse transcriptase (RT-p66/p51), together with a single molecule of transfer RNA^{lys}, which primes reverse transcription of the viral genome. The reverse transcriptase and integrase enzymes function at a post-entry step in the viral life cycle and mediate reverse transcription of the viral RNA genome into double stranded DNA and the integration of the double-stranded DNA into the host genome respectively. Viral proteins Vif and Nef are also thought to be closely associated with the nucleocapsid core.

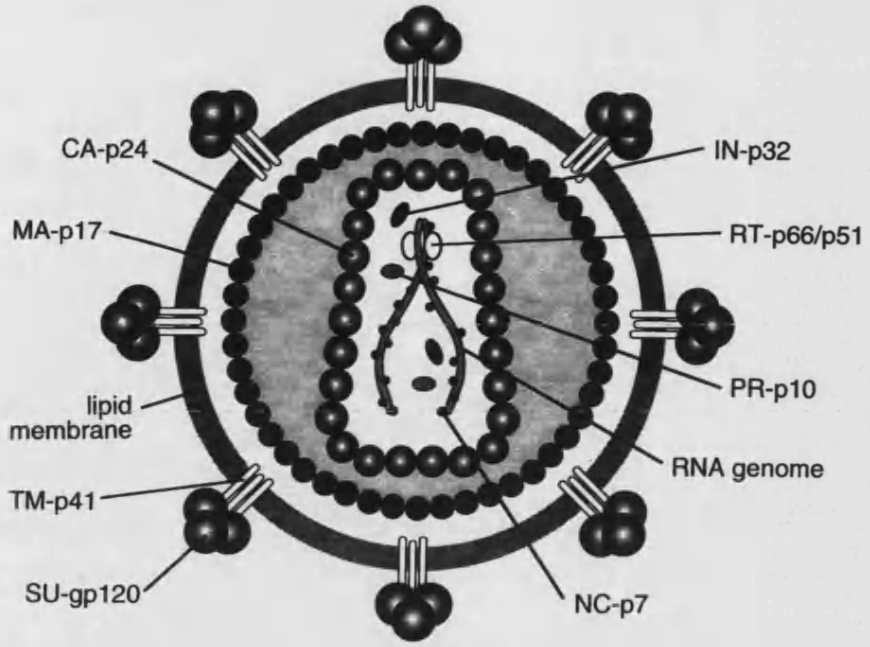


Figure 1.3 Schematic diagram showing the structural organisation of HIV (adapted from Stryer, 1988)

In accordance with a property more closely associated with a B type than a C type lentiviral particle the surface of the HIV virion is made up of virally encoded oligomeric envelope spikes (Chan *et al.,* 1997, Fass *et al.,* 1997, Weissenhorn *et al.,* 1997). In addition, because the lipid bilayer surrounding the virion particle is host-derived it also contains a number of host-cell proteins, notably MHC class I and class II molecules. The virally encoded glycoprotein structures are derived from a 160-kDa protein, gp160, which is co-translationally translocated into the endoplasmic reticulum (E.R.) by means of an N-terminal signal sequence. Inside the E.R. the signal peptide is removed and the protein is extensively glycosylated, specific disulphide bridges are formed and the protein forms trimers. In the Golgi apparatus, the gp160 protein is cleaved by cellular proteases into a surface (SU-gp120, HIV-1; SU-gp105, HIV-2) and transmembrane part (TM-gp41, HIV-1; TM-gp36, HIV-2) that remain associated through non-covalent interactions. The resulting glycoprotein is transported to the cell membrane and presented on the surface of the virion (Swanstrom & Wills, 1997). It is the SU component of gp160 which functions to adsorb the virion onto target cells via interactions with the cellular CD4 molecule and chemokine coreceptors.

The envelope glycoprotein is highly variable in amino acid sequence, comparison of *env* sequences taken from numerous isolates of HIV-1 has identified five hypervariable regions (V1-V5) interspersed with five conserved regions (C1-C5) for the gp120 subunit (Modrow *et al.,* 1987, Leonard *et al.,* 1990, Myers *et al.*, 1990). The conservation of the C1-C5 regions between isolates suggests that they are of biological significance; these regions form discontinuous structures important for interactions with the TM

ectodomain and viral receptors on the surface of the target cell. The first four variable regions form surface exposed loops with disulphide bonds at their bases (Leonard *et al.,* 1990). In addition to the envelope glycoprotein providing the virion with a mechanism for cellular entry, Env contains the major neutralising epitopes for neutralising antibodies produced during infection. Thus, the hypervariability and heavy glycosylation of Env modulates both the antigenicity and immunogenicity of the abundantly exposed protein (Profy *et al.*, 1990).

1.8 Viral regulatory and accessory proteins

In addition to the essential structural and enzymatic proteins encoded by human and simian immunodeficiency viruses, the HIV genome encodes a number of regulatory and accessory proteins. Regulatory proteins Tat (*trans*activator of *t*ranscription) and Rev (*r*egulator of *e*xpression of *v*irion proteins) are essential for viral growth and have been shown to regulate viral gene expression at the transcriptional and post-transcriptional levels, respectively. Accessory proteins Nef, Vif, Vpr (Vpx for HIV-2), and Vpu have also been shown to play a role in different steps of the virus life cycle, although none of these proteins are an absolute requirement for HIV replication.

Tat is synthesised at both early and late stages of the viral replication cycle and is not packaged into the virion. This protein binds to an RNA stem-loop element located within the viral LTR termed the *trans*activation *r*esponse (TAR) element, and functions to upregulate transcription of full-length viral transcripts (see section 1.11.1). The HIV-2 Tat protein is larger than HIV-1 Tat, 130 amino acids vs. 86 (Arya, 1993), but both function to activate transcription in a similar manner. HIV-1 Tat has been shown to activate transcription from the HIV-1 LTR and HIV-2 LTR with equal efficiency however, HIV-2 Tat *trans*activates expression from the HIV-1 LTR less efficiently than HIV-1 Tat. Further discussion of the structure, method of action and biological properties of HIV-2 Tat (Tat2) can be found in sections 1.10.4 and 1.11.

Rev is also a protein found at the early and late stages of the virus life cycle and in an analogous manner to Tat, targets an RNA stem-loop structure. This structure, termed the Rev Responsive Element (RRE), is located within the *env* gene. Rev has a central role in controlling the ratio of multiply spliced and unspliced RNAs in the cytoplasm (Chang & Sharp, 1989), it has been shown to require a cellular nuclear export protein called CRM-1 for its function (Fukuda *et al.*, 1997, Neville *et al.*, 1997). Multimerisation of Rev at the RRE promotes the nuclear export of RRE-containing unspliced or singly spliced RNA. Despite differences between HIV-1 and HIV-2/SIV Rev proteins, the mechanism of action of these proteins within their respective replication cycles appear to be the same.

Nef or *ne*gative *f*actor is a gene encoding a 27kDa myristolated cytoplasmic protein that is not essential for viral growth *in vitro*, but which is essential to the development of AIDS (Trono, 1995). The *nef* gene of HIV-1 and SIV_{cpz} extends from the 3' end of *env* into the U3 domain of the 3' LTR, whereas in HIV-2 and several strains of SIV the 5' end of *nef* overlaps the 3' end of *env* in a different translation frame.

Initially the Nef protein was thought to suppress HIV transcription by binding to a negative regulatory element (NRE), located in the U3 region of the LTR (Niederman *et al.*, 1989), however, these observations proved controversial (Hammes *et al.*, 1989, Kim *et al.*, 1989). Since then a wide variety of other functions have been associated with the Nef protein including enhancement of virion infectivity, downregulation of cell surface CD4 and class I major histocompatibility complex (MHC) (Greenberg *et al.*, 1998, Mangasarian *et al.*, 1999), upregulation of Fas-ligand (CD95L) expression on virus infected cells (Xu *et al.*, 1997, Xu *et al.*, 1999), and more recently, an anti-apoptotic effect of Nef has also been reported (Wolf *et al.*, 2001). However, it remains uncertain as to which *in vitro* function of Nef is most critical for the *in vivo* pathogenicity of primate lentiviruses. What is apparent is that many functions allow Nef to play an important role in the evasion of T cell recognition by HIV.

Defects in the *nef* gene and the function of the Nef protein have been linked to non-progressive HIV and SIV infection (Tobiume *et al.*, 2002, Tomiyama *et al.*, 2002). Rhesus monkeys inoculated experimentally with SIV containing deletions of *nef* have low viral loads, normal CD4 T-cell numbers, and are protected against challenge with virulent SIV (Kestler *et al.*, 1991, Daniel *et al.*, 1992). Similarly, numerous studies have documented the presence of *nef*-defective HIV-1 genomes in LTNPs (Deacon *et al.*, 1995, Kirchhoff *et al.*, 1995, Mariani *et al.*, 1996, Salvi *et al.*, 1998, Learmont *et al.*, 1999, Geyer *et al.*, 2001). Proviral Nef sequencing has shown that 10% of HIV-2 infected patients harbour truncated Nef proteins (Switzer *et al.*, 1998), higher than the prevalence seen with HIV-1 (Piguet & Trono, 1999), indicating that truncation of Nef may contribute to the slower disease progression rates observed in HIV-2 infections. More recently, a study performed by Padua and coworkers (Padua *et al.*, 2003) has identified a mutation within the tetra-proline motif of the HIV-2 Nef protein that is associated with an asymptomatic phenotype (Fisher's exact test $p = 0.026$) and low viral loads, however; the functional significance of this sequence change has yet to be determined. Despite these findings however, most of the *nef* alleles isolated from LTNPs of HIV-1 and HIV-2 infection are intact in the length of the coding region and in the tested biological function (Huang *et al.*, 1995a, Huang *et al.*, 1995b, Michael *et al.*, 1995, Catucci *et al.*, 2000), arguing against the role of *nef* in the establishment of long-term non-progression.

None of the remaining accessory proteins including Vif, Vpr, and Vpx (Vpu in HIV-1) are absolutely required for HIV replication; however, cell-type specific contributions to virus production and infectivity have been described (Trono, 1995, Cohen *et al.*, 1996, Madani & Kabat, 1998). Conservation of these sequences *in vivo* implies that the functions they perform are of some importance.

1.9. The replication cycle of HIV

The replication cycle of the simian and human immunodeficiency virus can be considered as a process that occurs in two phases, an early, and a late phase. Adsorption of the virus particle to a cell surface receptor, virus entry, reverse transcription of the viral genome and integration of proviral DNA into the host cell genome are all considered early phase events within the life cycle. Late phase processes involve viral gene expression from the integrated proviral template, virion assembly, and release of the progeny virions following maturation (Figure 1.4, pg 56).

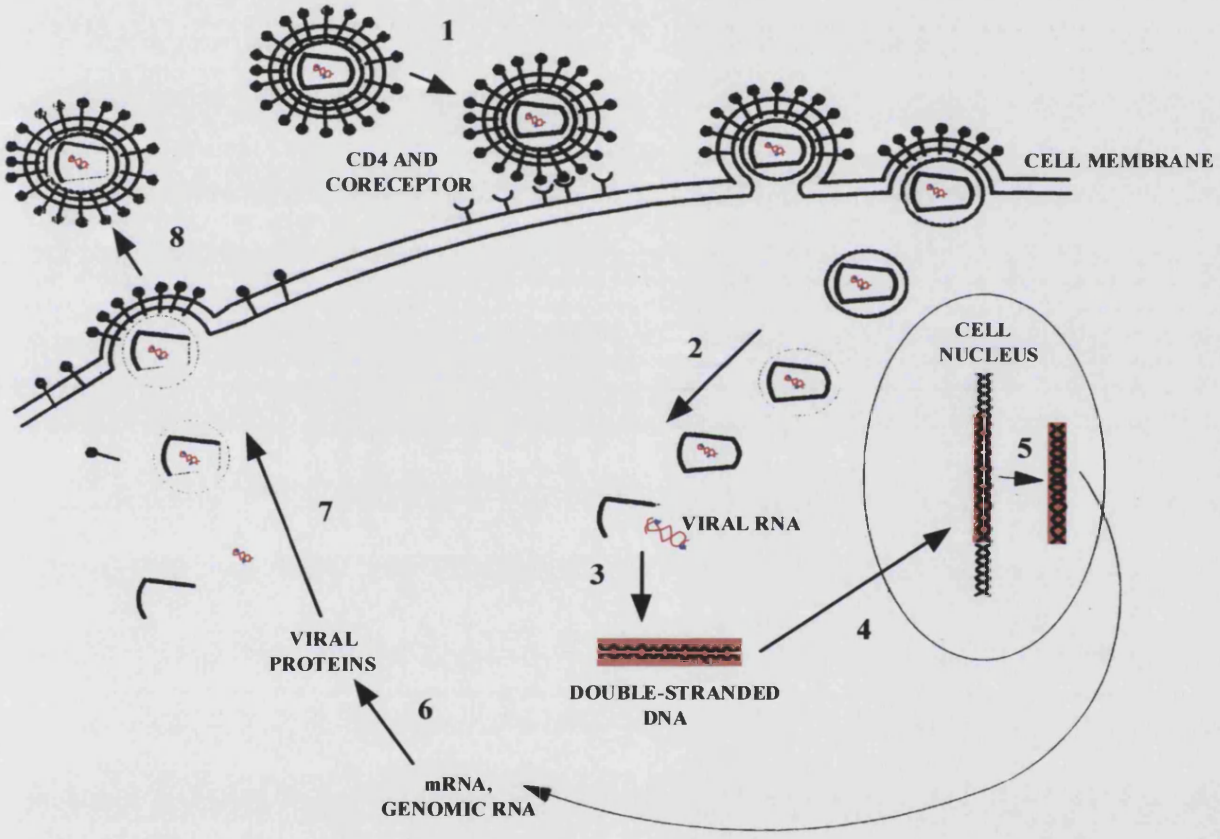


Figure 1.4 Generalised replication cycle of HIV. 1 – Viral attachment to specific target cell via interactions between viral envelope glycoproteins and cell surface CD4 and coreceptors. 2 – Fusion of viral and target cell membranes results in viral entry and uncoating. 3 – Reverse transcription of the viral RNA produces a double-stranded DNA. 4 – Integration of the dsDNA into the host cell genome resulting in a stably integrated provirus. 5 – Host cell factors transcribe the proviral DNA. 6 – Translation of viral mRNA produces the viral proteins. 7 – Viral proteins and full length viral RNA assemble into viral particles at the host cell membrane and acquire Env during budding. 8 – Gag and Gag-Pol polyproteins are cleaved by viral protease during or shortly after budding, generating mature infectious virions.

1.9.1 Viral attachment and entry

The attachment of HIV-1, HIV-2, and SIV to specific cell surface receptors is primarily mediated by interactions between the viral envelope glycoprotein and a 58kDa cellular glycolipid molecule called CD4. The CD4 antigen can be found located on the surface of T-helper lymphocytes, monocytes, macrophages and other cell types and was identified as the primary cellular receptor for HIV in 1984 following immunoprecipitation studies using the HIV envelope glycoprotein (Dalglish *et al.*, 1984, Klatzmann *et al.*, 1984a). This finding was later confirmed by experiments utilising human CD4 negative cells, non-permissive to HIV infection. Transfection of these cells with complementary DNA encoding human CD4 rendered them permissive for HIV-1 binding, syncitium induction, and infection (Maddon *et al.*, 1988). Epitope mapping of the CD4 receptor has revealed that HIV-1 gp120 binds to an immunoglobulin-like domain situated between amino acids 40-82 (Landau *et al.*, 1988, Peterson & Seed, 1988, Arthos *et al.*, 1989, Ryu *et al.*, 1990, Wang *et al.*, 1990).

Although CD4 is established as the primary receptor for HIV, a study in which a mouse cell expressing human CD4 was susceptible to HIV binding but resistant to infection indicated that the CD4 receptor was neither sufficient nor the sole means for viral attachment (Tersmette *et al.*, 1989, Dragic *et al.*, 1992, Broder *et al.*, 1993, Harrington *et al.*, 1993). This led to a search for secondary cellular receptors involved in the fusion and entry of HIV virions.

In 1996 Berger and co-workers identified a protein expressed from a human cDNA clone which when introduced into HIV-resistant mouse cells expressing human CD4 rendered them susceptible to HIV infection (Feng *et al.*, 1996). Further analyses showed this fusion-inducing molecule to be a coreceptor for T-cell tropic HIV-1 strains although it was unable to support entry of macrophage-tropic HIV-1 strains. This molecule, identified as Fusin/LESTR, but now termed CXCR-4, has been shown to be a member of the β-chemokine receptor family and has a natural ligand called stromal derived factor-1 (SDF-1). Following this discovery, several groups investigated the possibility of blocking HIV infection by using β-chemokines (Bleul *et al.*, 1996, Feng *et al.*, 1996, Oberlin *et al.*, 1996). During these studies a β-chemokine coreceptor, CCR-5, was found to mediate the cellular entry of macrophage-tropic strains of HIV-1. Taken together it has been established that in addition to CD4 both T-cell tropic and macrophage-tropic HIV strains utilize distinct but related β-chemokine coreceptors for cell entry. HIV-2 strains utilise CXCR-4 and CCR-5 in a similar manner to HIV-1 (Heredia *et al.*, 1997), however with the identification of several other chemokine coreceptors including CCR-1, CCR-2a, CCR-2b, CCR-3, CCR-4,

Bonzo/STRL33, and BOB/GPR15 (Choe *et al.*, 1996, Liao *et al.*, 1997), strains of HIV-2 have been shown to be, in general, more promiscuous than HIV-1 strains in their use of coreceptors (Guillon *et al.*, 1998, McKnight *et al.*, 1998, Owen *et al.*, 1998, Unutmaz *et al.*, 1998, Reeves *et al.*, 1999). With ever increasing numbers of coreceptor molecules being identified, in addition to reports of a CD4 independent strain of HIV-2 derived from HIV-2_{ROD} that can use CXCR-4 as its primary receptor (Reeves *et al.*, 1999), the envelope-cellular receptor mechanism of cell-entry grows increasingly complex.

Envelope viruses, such as HIV, enter cells by a mechanism that results in the direct fusion of viral and cellular plasma membranes. The precise details of this mechanism are still unknown although the process is known to be pH-independent. Binding of the envelope glycoprotein to CD4 is thought to trigger conformational changes in the surface (gp120/gp105) and transmembrane (gp41/gp36) regions of the envelope structure (Chan *et al.*, 1997, Weissenhorn *et al.*, 1997, Kwong *et al.*, 1998, Wyatt *et al.*, 1998). These changes are thought to potentiate and enhance the interaction between the viral envelope and the coreceptor molecule and transform the complex into what is known as a 'fusogenic state' (Lapham *et al.*, 1996, Wu *et al.*, 1996, Hill *et al.*, 1997). It is thought that structural changes within the leucine zipper domain of gp41 cause the exposure of a hydrophobic fusion peptide located at the N-terminus of gp41 (Moore *et al.*, 1993, Wild *et al.*, 1994), which inserts into the lipid bilayer of the cell plasma membrane inducing fusion of the viral and host cell membranes (Brasseur *et al.*, 1988).

1.9.2 Reverse transcription

Once inside the target cell the virus particle is partially uncoated and forms a large nucleoprotein complex. This complex is rapidly transported to the nucleus via interactions with host cell karyopherin proteins. Once inside the nucleus the heterodimeric viral enzyme reverse transcriptase (RT) catalyses the synthesis of a linear double-stranded DNA copy from the viral RNA genome. Reverse transcription of the dimeric single stranded viral RNA molecule is initiated from a host $tRNA^{lys}$ primer, which is hydrogen bonded to the primer binding site (PBS) located 100-200 nucleotides from the 5' end of the viral RNA template. The viral RNA has direct repeats at its ends called R segments, which vary in different strains of retrovirus from 10-80 nucleotides. Following the R segment at the 5' end of the virus is the U5 region of 80-100 nucleotides, whose name indicates that it is unique to the 5' end. Preceding the R segment at the 3' terminus is the U3 segment of 170-1350 nucleotides, which is unique to the 3' end. Each of these segments contains viral promoter and enhancer elements crucial to viral gene

expression and are collectively termed the viral long terminal repeat or LTR. Further discussion of the structure and biological properties of the HIV LTR can be found in sections 1.10, 1.11, and 1.12.

Since DNA synthesis begins at a site only 100-200 nucleotides from the 5' end of the minus RNA strand, the virus must utilise a mechanism called strand-transfer in order to generate a DNA copy of the entire RNA genome. The U5, R, and U3 regions of the viral LTR are involved in this process. The process of strand-transfer begins, immediately following $tRNA^{lys}$ binding to the minus strand PBS. Reverse transcriptase interacts with this primer to catalyse the synthesis of the first DNA strand (i.e., minus strand) and extends it to the 5' end of the viral RNA template, generating a short DNA sequence called minus strong-stop DNA. The Ribonuclease H (RNAse H) activity possessed by the reverse transcriptase heterodimer degrades the RNA in the DNA/RNA hybrid, effectively removing the R region at the 5' terminus of the RNA template (Varmus, 1988, Goff, 1990). Its removal exposes the newly synthesised strong-stop DNA sequence, which is complementary to the R region at the 3' end of the RNA strand (DeStefano *et al,* 1991). The complementary regions base pair with one another signifying the first strand-transfer event. Minus strand DNA synthesis continues on the viral RNA template through U3 and into the viral genome. The result of the switching event is to add a U3 segment to the 5' end of the minus strand. Generation of the plus strand of DNA involves a second transfer event whereby DNA is elongated to the end of the minus strand template and forms a plus strand strong-stop sequence. This sequence base pairs with the 3' end of the minus strand strong stop DNA, leaving RT free to complete the synthesis of the double stranded DNA copy of the viral genome. Circularisation of resulting double-stranded DNA copy is mediated by the LTR sequences at either end of the genome leading to the formation of the pre-integration proviral structure.

A high degree of nucleotide variation is introduced into the DNA copy of the viral genome during the process of reverse transcription because reverse transcriptase lacks a 3'-5' exonuclease proof reading function. Furthermore, because there are two RNA strands in each retroviral particle additional variation can be introduced by reverse transcriptase switching to the end of a different RNA molecule during the process of strand-transfer. Recombination events such as this can also occur in a cell that has been infected by more than one virion, variation can arise as a result of switching errors between different virion strands or copackaging errors. These processes lead to the generation of a diverse population of progeny virions, which can have altered viral properties of advantage or disadvantage to the replicative capacity of the population or its susceptibility to neutralisation by host immune responses (Hu & Temin, 1990).

1.9.3 Integration of viral DNA into cellular DNA

Once the reverse transcription of the viral genome is complete, the integrase enzyme binds to the double stranded viral DNA at the *att* site in the LTR. Integrase then catalyses the removal of two bases from the 3' ends of the viral DNA, resulting in the exposure of a highly conserved CA dinucleotide motif (Goff, 1992). The free hydroxyl groups' left as a consequence of this cleavage are covalently ligated to the 5' ends of IN-cleaved host-cell DNA by host cell enzymes to produce an integrated provirus (Bushman *et al.*, 1990, Craigie *et al.*, 1990, Katz *et al.*, 1990, Engelman *et al.*, 1991a).

The cellular DNA target site for integration of the HIV genome is not randomly selected. HIV-1 has been shown to preferentially insert into or near two repeated DNA elements in the human genome, L1 and *Alu*. The arrangement of the chromatin structure surrounding these sites is thought to determine the susceptibility to integration. In addition, the L1 and *Alu* elements have been found to share properties with retroviruses and have been subsequently termed retrotransposons.

1.9.4 Viral protein expression and regulation of expression

Once integrated into host cell DNA, the provirus relies on the complex interplay between *cis*-acting viral DNA/RNA sequences, viral *trans*activator proteins, and host-cell transcription factors to co-ordinate the expression of HIV viral proteins (see section 1.10). The sequence elements that interact with the viral and host-cell factors are located in the promoter/enhancer region of the viral genome termed the LTR.

Since the integrated provirus contains a 5' and a 3' LTR that are identical in sequence, the question arises as to which LTR functions as the transcriptional promoter. Studies undertaken to investigate the promoter activity of the 5' and 3' HIV-1 LTRs demonstrated that when placed upstream of a proto-oncogene both LTRs were functional, although absolute levels of transcription directed by the 3' LTR of HIV were reduced compared to the 5' HIV LTR promoter (Klaver *et al.*, 1994). In addition, the transcriptional activity of the 3' LTR was dependent on the deletion or inactivation of the 5' LTR. These results confirmed similar findings investigating the LTR activities of avian retroviruses (Fung *et al.*, 1981, Neel *et al.*, 1981, Payne *et al.*, 1981). Since then, several groups have demonstrated that HIV gene expression of the integrated genome *in vivo* is directed exclusively by the 5' LTR. The elongation complexes produced at the 5' LTR have the ability to prevent or disrupt the assembly of pre-initiation complexes at the downstream 3' LTR (Cullen *et al.*, 1984, Klaver *et al.*, 1994), in a mechanism termed transcriptional interference or promoter occlusion (Adhya & Gottesman, 1982).

Transcription of integrated proviral DNA is directed by the cellular RNA polymerase II (RNAPII) enzyme and its transcription efficiency is dramatically enhanced following the expression of the viral *trans*activator protein Tat. The mechanism of regulated viral transcription is discussed in detail in section 1.10 and 1.11.1.

Like most mammalian mRNAs, HIV transcripts are capped at the 5' end and polyadenylated at the 3' end. Capping involves the enzymatic addition of a cap group (m7G5'pppN) to the most 5' nucleotide of each nascent viral transcript. The addition of this group enables the components of the cellular translational machinery to recognise the viral RNA (Lewis & Izaurralde, 1997). Host cell enzymes also carry out polyadenylation of the viral RNA. An AAUAAA (poly A) signal located within the HIV LTR directs polyadenylation of the 3' end of the RNA. Because this signal is found within the LTR sequence, it is present at both the 5' and 3' end of the transcript. Polyadenylation at the 5' site would yield a short, non-coding RNA, reducing the amount of full-length viral RNA. In order to prevent this from occurring HIV has developed a regulatory mechanism involving the occlusion of the polyA signal in a stable RNA structure (Klasens *et al.*, 1998, Das *et al.*, 1999). This suboptimal setting allows both complete suppression of the 5' signal due to 5' specific inhibitory elements and full activity of the 3' signal due to 3' specific enhancer elements (Das *et al.*, 1999).

Early after infection, multiply spliced transcripts are produced predominantly, which encode the regulatory Tat, Rev, and Nef proteins. Expression of the viral Rev protein shifts the balance from multiply spliced transcripts toward the appearance in the cytoplasm of singly spliced and unspliced RNAs that encode the structural Gag, Pol and Env proteins. Rev is thought to influence the frequency of splicing by dissociating spliceosome components from viral transcripts (Chang & Sharp, 1989), thus accumulation of Rev will dissociate more spliceosomes resulting in fewer splicing events in new viral transcripts. The RRE of HIV-2 and SIV are substantially different in secondary structure to the HIV-1 RRE. HIV-1 Rev proteins can induce HIV-2 protein expression by interacting with HIV-2 RRE, however HIV-2 Rev fails to activate HIV-1 protein expression (Dillon *et al.*, 1991).

1.9.5 Virion assembly, budding and maturation

Assembly of the virion particle components begins with the accumulation of p55 Gag and p160 Gag-Pol polypeptide complexes at the inner surface of the host-cell plasma membrane (Gottlinger *et al.*, 1989, Bryant *et al.*, 1991, Zhou *et al.*, 1994). A *cis*-acting packaging element located near the 5' end of the full-length viral transcript directs the association of the viral genome with these polypeptide complexes by

recognising the NC domain in the Gag polypeptide. In addition to Gag polypeptide accumulation, envelope glycoprotein oligomers are also targeted to the plasma membrane and, once inserted, interact with the MA domain of p55 Gag via the cytoplasmic domain of the TM subunit.

One of the most important events to take place during virion assembly involves the proteolytic cleavage of the Gag, Gag-Pol, and Env-gp polypeptide precursors. The resulting mature nucleocapsid is composed of fully processed *gag* (MA, CA, NC, p6) and *pol* gene products (PR, RT, RNase H, and IN) as well as two molecules of viral single-stranded RNA genome. The final stage of the virus life cycle is achieved once the nucleocapsid core buds through the host-cell membrane, producing a virion particle with a nucleocapsid surrounded by a lipid bilayer membrane containing envelope glycoprotein spikes and other host-cell membrane proteins.

1.10 The control and regulation of HIV-2 gene expression

Viruses are obligatory intracellular parasites and are optimally adapted to use the host cell machinery for the production of new virus particles. In order to achieve this HIV-2 possesses a promoter region that contains *cis*-acting regulatory sequences that are recognised by host cell transcription factors. This enables transcription of HIV-2 encoded genes to be executed by the same enzyme system that expresses cellular genes, and primarily involves cellular enzyme RNA polymerase II (RNAPII). Since the activities of several of the cellular transcription factors that recognise elements within the HIV-2 promoter are altered by parameters such as cellular activation or differentiation, HIV-2 gene expression is also, in part, controlled by alterations in cellular signal transduction pathways. In order to put the sequence elements of the HIV-2 LTR into the context of the eukaryotic transcription process when described in the next section of this chapter, a brief overview of eukaryotic transcription is given below.

1.10.1 Eukaryotic gene expression

The initiation of transcription from a eukaryotic and thus the integrated HIV-2 promoter can be divided into three major steps: chromatin remodelling, transcription factor recruitment, and elongation.

1.10.1.1 Chromatin remodelling

Within the cell nucleus, DNA is packaged with nuclear histone proteins into a structure known as chromatin. Within this complex 200bp stretches of DNA become wrapped around a core of histone proteins to form the basic subunits of chromatin and are termed nucleosomes. This organisation exerts a repressive effect upon gene expression and determines that most DNA sequences are structurally inaccessible and functionally inactive. As a result, the efficiency of transcription initiation is determined, in part, by the accessibility of the promoter embedded within the chromatin structure. In order for transcription from a eukaryotic or an HIV promoter to be initiated remodelling of the chromatin structure must occur (Verdin, 1991, Verdin *et al.,* 1993, Van Lint *et al.,* 1996, Widlak *et al.,* 1997, Benkirane *et al.,* 1998, Kadonaga, 1998, Marzio *et al.,* 1998, Deng *et al.,* 2001, He *et al.,* 2002). Structural rearrangement of chromatin can be brought about during the process of transcription by histone acetylation (Struhl, 1998, Berger, 2002, Eberharten & Becker, 2002). Histone acetylatransferases (HATs) have been shown to modify histones by covalently linking acetyl moieties to the free amino groups of the amino acids that make up the proteins, resulting in the removal of the positive charge that resides on the NH_3^+ side chain (Brownell & Allis, 1996). It has been speculated that the reduction of positive charges on histone

proteins might lower their affinity for DNA, resulting in a relaxation of the DNA-protein structure thereby facilitating the binding of sequence-specific transcription factors to the promoter DNA. Several multisubunit ATP-utilising enzymes which can remodel chromatin structure *in vitro* have been identified, including mammalian HAT and the CBP (CREB-binding protein)/p300 family of coactivator proteins and their associated cofactor (P/CAF) (Chan & La Thangue, 2001, Berger, 2002). Moreover, recent analyses have indicated that, in one of its many roles in the upregulation of transcription, Tat functions to recruit a HAT to the promoter (see section 1.11.1ii).

1.10.1.2 Transcription factor recruitment for PIC assembly

Once chromatin remodelling has occurred the process of transcription factor recruitment and pre-initiation complex (PIC) formation begins. The most recent and widely supported model for eukaryotic and thus HIV PIC assembly consists of two steps (Berk, 1999, Buratowski, 2000, Gill, 2001, Butler & Kadonaga, 2002).

The first step involves the binding of general transcription factors TFIID and TFIIA to the promoter (Coleman *et al.,* 1999, Liu *et al.,* 1999) (see Figure 1.5A, pg 68). TFIID is a multisubunit complex that has been shown to be only one of two general transcription factors to have sequence specific DNA-binding activity. Importantly, TFIID recognises the TATA box located within the eukaryotic or HIV-2 promoter as its DNA target. The TFIID complex is composed of the TATA-binding protein (TBP), a small protein of 30 kDa, and a number of tightly complexed TBP-associated factors (TAFs) (Burley & Roeder, 1996). In recognising the TATA sequence the TBP subunit of TFIID binds to the minor groove of the promoter DNA, forming a 'saddle' around the double helix (Kim *et al.,* 1993). In effect, the inner surface of TBP binds to the DNA while the larger outer surface is available for interaction with other proteins. Crystal structures of eukaryotic TBP bound to short TATA-box-containing promoter fragments have revealed that binding of TBP to the promoter highly distorts the DNA, causing it to bend 75-80° towards the major groove (Kim *et al.,* 1993, Coulombe & Burton, 1999). This process is thought to promote the second step of PIC assembly, which involves the binding of a holoenzyme complex containing RNA polymerase II (RNAPII) and the remaining GTFs required to initiate transcription (Koleske *et al,* 1995). The TFIID complex as a whole typically protects the region from -45 to –10 in a eukaryotic promoter, encompassing the TATA box and its surrounding nucleotides. General transcription factor TFIIA has been found to stimulate the rate and extent of TFIID binding and acts to stabilize the DNA-TFIID complex (Coleman *et al.,* 1999, Liu *et al.,* 1999). After joining the complex, TFIIA extends

protection of the TFIID region further upstream (Figure 1.5A, pg 68), although it is not clear if this factor remains associated when the PIC is fully assembled.

The second step of PIC assembly involves the recruitment to the DNA-TFIID-TFIIA complex of the holoenzyme containing RNAPII and a number of associated GTFs (Figure 1.5B, pg 68). While the composition of the definitive holoenzyme has yet to be determined, general transcription factors TFIIB, TFIIE, TFIIF, and TFIIH have been shown to play crucial roles (Orphanides *et al.*, 1996). TFIIB, the only other general transcription factor to possess sequence specific DNA-binding activity, is thought to interact with TFIID and the DNA major groove, binding to a bipartite sequence termed the "IIB recognition element" (BRE) located nine base pairs upstream, (and seven base pairs downstream) of the eukaryotic TATA element (Lagrange *et al.*, 1998, Littlefield *et al.*, 1999). Binding of TFIIB to the BRE has been shown to stabilize the bend in the DNA and determine the orientation of TBP binding to the TATA element, in effect setting the direction of the DNA distortion (Nikolov *et al.*, 1995, Lagrange *et al.*, 1996). TFIIF is a GTF composed of two subunits, RNA polymerase-associating protein 74 (RAP74) and RAP30, and has been shown to be an absolute requirement of eukaryotic transcription both *in vivo* and *in vitro* (Lei *et al.*, 1999, Yan *et al.*, 1999). The RAP30 subunit of TFIIF binds directly to RNAPII and to TFIIB and mediates the delivery of the RNAPII-containing holoenzyme to the PIC. General transcription factors TFIIE and TFIIH have also been identified within the holoenzyme complex. The presence of TFIIH within the PIC is particularly significant in transcription initiation as it is the only GTF to possess enzymatic properties, containing two helicase subunits, XPD and XPB, and a cyclin dependent protein kinase subunit, which acts to phosphorylate the CTD of RNAPII (Tirode *et al.*, 1999) (see section 1.10.1iii).

Binding of the holoenzyme and its associated general transcription factors cause further distortion of the DNA resulting in slight strand separation. It has been proposed that the separation of the DNA strands is enough to allow the XPB helicase subunit of TFIIH access to single-stranded DNA to catalyse open-complex formation in the presence of ATP (Coulombe & Burton, 1999). The process of DNA unwinding allows RNAPII to read, and transcribe from, the coding strand of the DNA (Tirode *et al.*, 1999).

Significantly, the complex interactions between TFIID, RNAPII, TFIIB, and the TATA and BRE elements (that result in conformational changes in DNA structure) determine the location of the open complex and thus the transcriptional start site position of +1 (Kim *et al.*, 1997).

The rate at which general transcription factors are recruited to the promoter, and thus the rate at which transcription initiation can occur, is dependent on activator proteins that usually have binding sites upstream of the TATA box element. Activator proteins exert their influence upon the rate of transcription by either binding to and recruiting one or more of the general transcription factors described above, or in binding to the promoter in an upstream position they enhance the binding stability of recruited general transcription factors (Blau *et al.,* 1996, Gill, 2001). The properties of such proteins and their binding sites within the HIV-2 promoter will be discussed in detail in section 1.10.2.

1.10.1.3 Promoter clearance and elongation

The third and final stage in the process of transcription initiation is termed promoter clearance. The mechanisms involved in this process are still not fully understood, however, they result in the transition of the PIC to an elongation complex.

Initial studies into the third stage of transcription initiation had documented a rate-limiting step *in vivo.* Following the initiation of RNA synthesis, PICs were shown to stall and accumulate 20 to 40 nucleotides downstream of the eukaryotic and HIV promoter (Kao *et al.,* 1987, Kessler & Mathews, 1992). How these PICs were transformed into complexes generating full-length transcripts was undetermined until extensive analyses were performed upon the RNAPII enzyme within the PICs themselves. The data generated by these analyses demonstrated that the release of the stalled pre-initiation complexes and their transformation into elongating complexes required the phosphorylation of the CTD of the largest RNAPII subunit (Marciniak & Sharp, 1991, Marshall, *et al.,* 1996, Okamoto, *et al.,* 1996).

The CTD of RNAPII contains multiple heptapeptide tandem repeats of Tyr-Ser-Pro-Thr-Ser-Pro-Ser. Within preinitiation complexes (Laybourn & Dahmus, 1989) and in early elongation complexes *in vitro* (Marshall, *et al.,* 1996) the serine residues within the multiple tandem repeats of the CTD were found to be hypophosphorylated (RNAPIIa). During the process of promoter clearance and productive elongation however, the serine residues within the CTD were found to be exclusively hyperphosphorylated (RNAPIIo).

The protein kinase responsible for the initial phosphorylation of the CTD of RNAPII during promoter clearance has been shown to be the cyclin-dependent kinase subunit (cdk7) of general transcription factor TFIIH, in association with cyclin H and MAT1 (collectively known as cyclin-dependent kinase activating kinase CAK) (Goodrich & Tjian, 1994, Tirode *et al.,* 1999) (Figure 1.5C, pg 68). The allosteric changes in the conformation of RNAPII induced by CAK phosphorylation of the CTD are thought to release the

PIC from the promoter by disrupting the contacts between the CTD and other transcription factors such as TFIID which do not remain with the elongation complex (Zawel, *et al.,* 1995) (Figure 1.5D, pg 68). When the melted region or the RNA transcript reaches a certain critical length RNAPII releases its contacts with the promoter DNA and associated general transcription factors, thus entering the elongation phase of transcription. In support of this model, TBP cannot associate with RNAPIIo (Usheva, *et al.,* 1992).

More recent analyses have indicated that while the CAK complex of TFIIH may facilitate promoter clearance, it does not appear to be the kinase required during the elongation process. Identification of human positive transcription elongation factor b (P-TEFb) has lead to the proposal that the kinase subunit of this factor is responsible for the hyperphosphorylation of the CTD, and subsequent elongation of the RNA transcript (Peng *et al.,* 1998). The kinase subunit of P-TEFb is composed of cyclin-dependent kinase 9 (cdk9), and cyclin T1 (CycT1), although a small proportion of the factor (<10%) is associated with either cyclin T2a or cyclin T2b (Peng *et al.,* 1998). This factor has been shown to hyperphosphorylate the CTD and is sensitive the kinase inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), a well-known inhibitor of transcription elongation (Marshall *et al.,* 1996). The large conformational changes induced by RNAPII CTD hyperphosphorylation permit RNAPII to interact with a wide variety of elongation factors as well as with proteins involved in mRNA processing (Hirose & Manley, 2000).

Notably, several studies have recently demonstrated that the HIV *trans*activator protein Tat mediates its enhancement of gene expression by recruiting P-TEFb to the promoter, effectively increasing the rate of elongation (Alonso *et al.,* 1994, Mancebo *et al.,* 1997, Zhu *et al.,* 1997, Fujinaga *et al.,* 1998, Wimmer *et al.,* 1999), this will be further discussed in section 1.11.1.

Transcription termination is coupled to 3' end processing, which involves cleavage of the transcript and poly (A) addition (Guntaka *et al.,* 1993). Recent reports indicate that transcription by RNAPII and pre-mRNA processing are linked in a number of ways, since essential subunits for mRNA capping, splicing and polyadenylation associate with the RNAPII enzyme during transcription (Neugebauer & Roth, 1997).

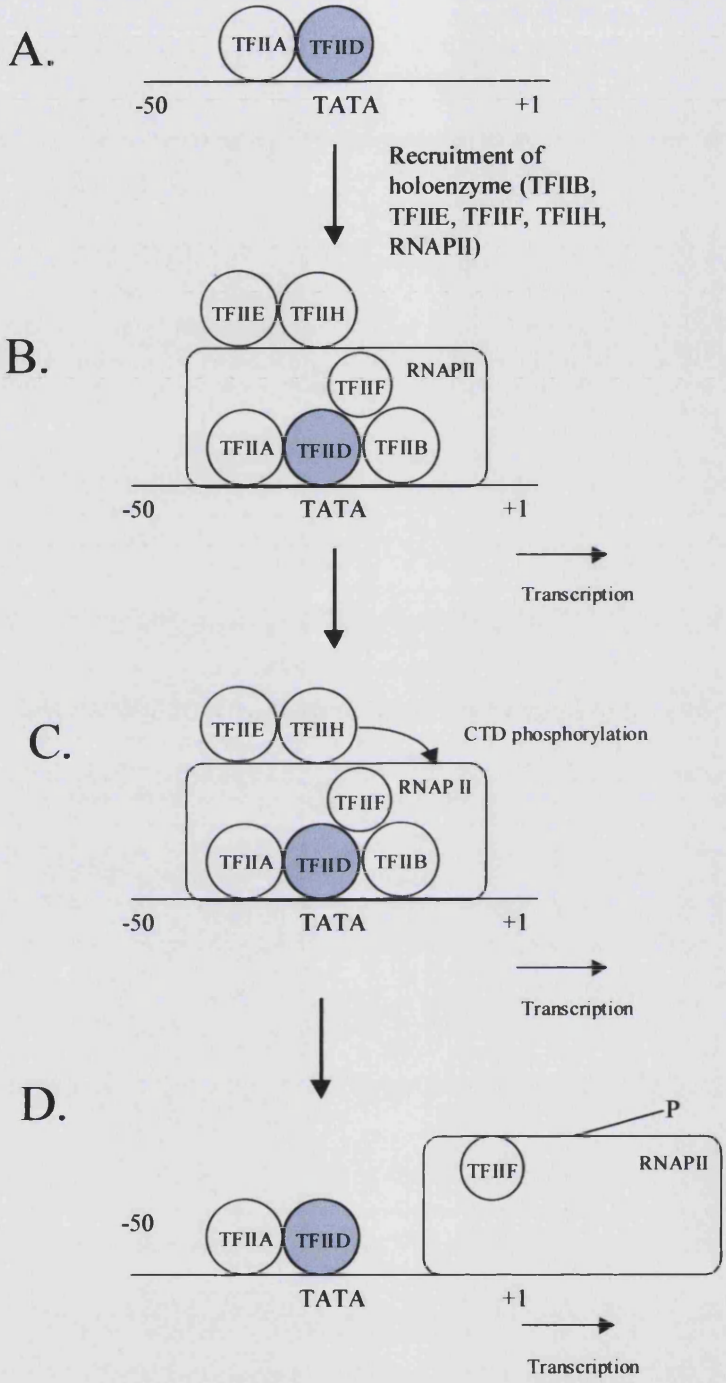


Figure 1.5 Diagram of the processes involved in the initiation of transcription from a eukaryotic (or an integrated HIV-2) promoter. (A). Initially the PIC is assembled, general transcription factor TFIIH binds to the TATA box together with TFIIA. (B). The second step of PIC assembly involves recruitment to the promoter of a holoenzyme containing RNAPII and the remaining GTFs required to initiate transcription. (C). The cyclin-dependent kinase subunit (cdk7) of GTF TFIIH catalyses the phosphorylation of the CTD of RNAPII, leading to a progression of conformational changes in DNA, RNAPII, and the GTFs that result in the formation of an open complex capable of forming the first phosphodiester bond. (D). Phosphorylation of the CTD of RNAPII has converted the pre-initiation complex into to an elongation complex. Contacts between the CTD and other transcription factors such as TFIID, which do not remain with the elongation complex, are disrupted and the elongation complex is free to move down the gene producing the RNA transcript.

1.10.2 The structure of the HIV-2 Long Terminal Repeat (LTR).

As described in section 1.10.1, the complex array of putative regulatory elements that interact with the viral and cellular factors to cause the expression of all HIV-2 encoded genes, are located within the region of the HIV-2 genome termed the LTR. Complementary experiments and nucleotide sequence analyses of the original HIV-2 isolate (ROD) have revealed that the HIV-2 LTR is a prototypic enhancer-promoter unit, 854bp in length (Guyader *et al.,* 1987). Further analyses have identified three distinct internal domains within the LTR termed U3, R, and U5 (Figure 1.6A, pg 71) (Guyader *et al.,* 1987, Arya & Gallo, 1988).

The U3 domain of the HIV-2 LTR has been mapped between nucleotides –556 to –1 (Figure 1.6A, pg 71), relative to the start site of transcription, and contains most of the transcription factor binding elements required to direct HIV-2 gene expression (Guyader *et al.,* 1987, Tong-Starksen *et al.,* 1990). This domain can be further trisected into core, enhancer, and modulatory regions, corresponding to the nature of the binding elements located within them (see section 1.10.3).

The R domain of the HIV-2 LTR has been shown to contain sequences required for the transcriptional response to the viral *trans*activator protein Tat. This region, mapped to nucleotides +1 to +173, relative to the start site of transcription (Figure 1.6A, pg 71), encodes an RNA sequence element that forms the *t*rans*a*ctivation *r*esponse element (TAR), a structure only found in human and simian immunodeficiency viruses (Guyader *et al.,* 1987, Tong-Starksen *et al.,* 1990) (see section 1.10.4).

The U5 domain of the HIV-2 LTR is contained between nucleotides +174 to +299 (Figure 1.6A, pg 71), and may encode sequences that negatively regulate HIV-2 transcription (Guyader *et al.,* 1987, Tong-Starksen *et al.,* 1990, Arya, 1991) (see section 1.10.5).

Computer alignments of HIV LTR sequences have shown that all of the elements of the HIV-2 LTR are larger than the corresponding elements within the HIV-1 LTR (respectively (U3) 456 HIV-1; 556 HIV-2; (R) 97; HIV-1; 173; HIV-2; and (U5) 82; HIV-1; 120; HIV-2) (Guyader *et al* 1987). The entire HIV-2 LTR shares only 40% sequence homology with the HIV-1 LTR, although, some core transcriptional elements display 50% sequence similarity (Tong-Starksen *et al.,* 1990). As with the rest of the genome the HIV-2 LTR shares much greater sequence similarity with the SIV LTR (Chakrabarti *et al.,* 1987, Franchini *et al.,* 1987). In particular, at least three modulatory transcription factor binding sites within the U3 domain of the HIV-2 and SIV LTRs share 75% sequence homology and are distinct from any element

located in the modulatory region of the HIV-1 LTR (Hirsch *et al.,* 1989b, Renjifo *et al.,* 1990, Hannibal *et al.,* 1994).

The location and function of the *cis*-and *trans*-acting sequences located within the U3, R and U5 domains of the HIV-2 LTR have been determined in studies using gel retardation and deoxyribonuclease (DNase) I footprinting assays, in addition to plasmids containing LTR promoters linked to reporter genes such as bacterial chloramphenicol acetyl transferase or firefly luciferase. Site-directed mutagenesis of the binding elements has established whether their effect on gene expression is positive or negative and has also revealed synergyistic interactions between the transcription factors that bind them. Several studies have extended this approach by constructing HIV proviruses that contain mutations in potential regulatory regions. The growth properties of these mutant viruses have been compared to those of wild-type virus. Identification of the cellular factors that bind to the regulatory elements has been achieved using oligonucleotide affinity chromatography, λ gt11 expression cloning, and somatic cell genetics.

The next section of this chapter will describe the sequence elements and transcription factors shown to play a role in gene expression directed by the HIV-2 LTR. Data concerning the function of the HIV-2 LTR elements and their binding factors are limited in comparison to HIV-1, although the function of several of the core binding sites have been shown to be analogous (Arya & Gallo, 1988, Tong-Starksen *et al.,* 1990). Wherever possible data derived from HIV-2 LTR analyses has been described, however, some details included have been derived from analyses carried out on the HIV-1 LTR.

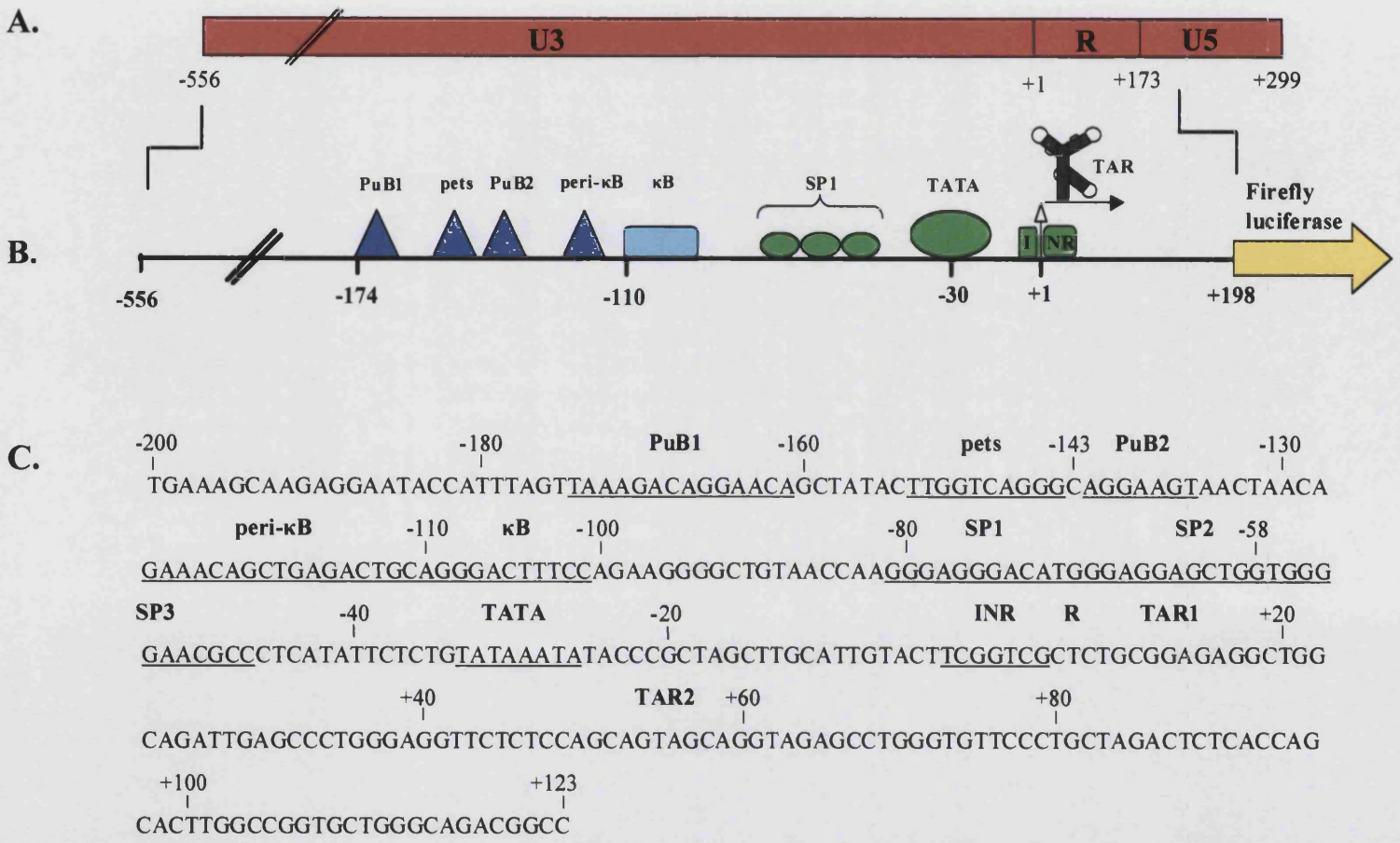


Figure 1.6 Schematic representation of the HIV-2 $pROD_{10}$ Long Terminal Repeat (LTR) (A) The domain structure of the HIV-2 LTR. The boundaries and sizes of the putative U3, R and U5 domains are shown (Guyader *et al* 1987). (B) Diagram of the cloned region of the HIV-2 LTR. Relevant sites within the HIV-2 LTR are identified. Those sites located within the core promoter region of U3 are shown in green, those within the enhancer region of U3 are shown in turquoise, and those within the modulatory region of U3 are shown in blue. (C) Sequence of the region of the HIV-2 LTR (ROD) from nucleotides –200 to +123. This region of the HIV-2 LTR contains the majority of the functionally relevant transcription factor binding elements identified thus far; the sequences of these elements are underlined.

1.10.3 The structure and function of the binding elements within the U3 domain of the HIV-2 LTR.

1.10.3.1 The core region of U3

The transcription factor binding sites defined within the core region of U3 have been shown to be essential to both basal and Tat-induced HIV gene expression (Guyader *et al.,* 1987, Arya & Gallo, 1988, Tong-Starksen *et al.,* 1990, Arya, 1993). In addition, all are analogous to the fundamental core sequence elements usually found within a generic eukaryotic/mammalian promoter. Taken together the elements within this region constitute the shortest sequence at which RNAPII can initiate viral gene expression (Figure 1.6B and C, pg 71).

1.10.3.1a.The TATA box and its surrounding nucleotides

As in the mammalian promoter, of central importance to the HIV-2 promoter is the highly conserved A-T rich TATA box motif that is located 30bp upstream from the transcriptional start site (Figure 1.6B and C, pg 71). Extensive analyses of the HIV-1 and HIV-2 LTR have ascertained that the TATA element is critical for both basal and Tat-induced gene expression (Guyader *et al.,* 1987, Arya & Gallo, 1988, Jakobovits *et al.,* 1988, Jones *et al.,* 1988, Garcia *et al.,* 1989, Tong-Starksen *et al.,* 1990, Berkhout & Jeang, 1992, Rittner *et al.,* 1995). The explanation for the fundamental influence that this 8bp motif has on both basal and activated levels of transcription lies in it acting as the initial DNA target site for the TATA-binding protein (TBP), the 38-kDa component of general transcription factor TFIID (Berkhout & Jeang, 1992, Gaynor, 1992, Ou *et al.,* 1994, Coulombe & Burton, 1999, Patikoglou *et al.,* 1999, Ranish *et al.,* 1999). As described in section 1.10.1, the binding of the complex of cellular factors that constitute TFIID is a prerequisite to the progressive assembly of the pre-initiation complex (PIC), the aggregation of general transcription factors and RNA polymerase II required for the initiation of transcription.

By acting as the binding site for TFIID, the HIV-2 TATA element defines the window within which transcription initiation and open complex formation can occur. In determining the location of the open complex, the interactions between, RNAPII, the general transcription factors and their binding elements define the actual position of the transcriptional start site position +1 (Kim *et al.,* 1997, Coulombe & Burton, 1999).

In vitro and *in vivo* transcription analyses of the HIV-1 and HIV-2 LTR indicate that mutation of the TATA element dramatically decreases HIV gene expression (Arya & Gallo, 1988, Jakobovits *et al.,* 1988,

Garcia *et al.*, 1989, Tong-Starksen *et al.*, 1990, Berkhout & Jeang, 1992, Rittner *et al.*, 1995). Garcia *et al* demonstrated that disruption of the HIV-1 TATA sequence caused loss of DNAse protection between nucleotides -42 - -4, the region to which TFIID binds, and showed that this loss was accompanied by marked decreases in Tat-induced transcriptional activation and HIV specific mRNA (Garcia *et al.*, 1989). Similarly Jakobovits *et al.*, showed severely impaired transcription rates from both HIV-1 and HIV-2 LTRs with mutations in the TATA element (Jakobovits *et al.*, 1988). Moreover, proviral constructs containing the TATA sequence mutations investigated by Garcia *et al* not only directed severely impaired levels of HIV gene expression but also failed to generate progeny virus (Harrich *et al.*, 1990).

Many of the studies that have characterised the effects of mutations within the HIV-1 TATA element have also revealed that sequences immediately flanking this motif are important in modulating HIV gene expression and may play a role in HIV TATA element function. In a study carried out by Berkhout *et al* thirty nucleotides immediately flanking the HIV-1 TATA motif were replaced with the corresponding sequences from the HTLV promoter (Berkhout & Jeang, 1992). The HTLV/HIV hybrid promoter did not work nearly as well as the wild type HIV promoter. In addition, transcription analyses carried out by several groups have demonstrated that site-directed mutagenesis of sequences immediately 5' and 3' to the HIV-1 TATA box results in marked decreases in both basal and Tat-induced gene expression (Garcia *et al.*, 1989, Berkout & Jeang, 1992, Ou *et al.*, 1994, Rittner *et al.*, 1995). In eukaryotic transcription, this region of the promoter provides the recognition sequences for the general transcription factor TFIIB, a factor that plays a central role in PIC formation and stability (Lagrange *et al.*, 1997). Changes to the nucleotides 7bp upstream and 9bp downstream of the TATA box which encompass the eukaryotic TFIIB recognition element (BRE) have been shown to result in reduced levels of transcription initiation as a consequence of decreased TFIIB binding. While a consensus BRE has not been identified in the HIV-1 or HIV-2 LTRs, competition analyses suggest that the region flanking the HIV TATA box may contain multiple recognition sequences for distinct cellular proteins that are required for stabilization of, or interaction with, factors which bind to the TATA enhancer region (Garcia *et al.*, 1989, Berkout & Jeang, 1992, Ou *et al.*, 1994).

1.10.3.1b. The HIV-2 Initiator element

Originally, the initiator (Inr) element was defined as a 17 bp motif that surrounded the RNA initiation site in TATA-less promoters (Smale & Baltimore, 1989) and contained all the information necessary for determining specific initiation of transcription by RNA polymerase II. Since then, several studies have

reportedly found similar elements within mammalian promoters, leading to the establishment of a loose consensus sequence for mammalian Inr of $YCA^{+1}NTYY$ where Y is a pyrimidine and N is any nucleotide (Javahery *et al.,* 1994).

A region sharing this loose consensus sequence and which appears necessary for promoter activity has been identified within the start site region of the HIV-2 LTR. Extensive deletion and substitution analyses have revealed that significant reductions in transcriptional activity are observed if the sequences between residues -6 and +30 of the HIV-1 and the HIV-2 LTR are mutated (Jones *et al.,* 1988, Rittner *et al.*, 1995). More detailed analyses of this region has revealed that residues +11 to +20 have no real effect upon promoter strength indicating that the HIV Inr may be a bipartite initiator element or two elements positioned between residues -6/+4 and +21/+30 which act synergistically to increase promoter activity (Zenzie-Gregory *et al.,* 1993, Rittner *et al.,* 1995).

The precise mechanism by which the Inr element contributes to promoter activity is as yet undetermined although it appears that the HIV Inr requires precise positioning relative to the TATA box in order for it to influence the rate of transcription initiation. A study investigating the significance of core promoter element positioning in the HIV-1 LTR demonstrated a 3-4 fold reduction in transcription initiation rates when the TATA box and HIV-1 Inr were separated by 15 and 22 nucleotides (Rittner *et al.,* 1995). No comparative studies exist for the HIV-2 LTR although in accordance with the function of other analogous elements, similar effects could be expected.

Speculation as to the identity of potential Inr-binding proteins (ITF) has predicted that the Inr might be required for transcription initiation based on its participation in the co-operative binding of the general transcription factor TFIID along with the TATA box. Binding studies have however, failed to prove that TFIID or any of its subunits bind to any Inr element. More recently, a study carried out by Roy *et al* has shown that the HIV-1 initiator element could be bound by TFII-I, a 120-kDa general transcription factor known to be involved in transcription initiation (Roy *et al.,* 1991). It has been suggested that interaction of the initiator binding protein with a member or members of the pre-initiation complex may in some way facilitate complex formation. Thus in binding to the HIV Inr TFII-I could potentially interact with TFIID, the central transcription factor of the PIC, and influence transcription initiation via a stabilising or recruitment effect. This would explain the functional requirement of the HIV Inr to be in close proximity to the TATA box. Comprehensive analysis of the HIV-1 and HIV-2 Inr and the factors which bind to them have yet to be completed.

1.10.3.1c. SP1-binding sites

It has been widely established that the PIC containing the GTFs and RNAPII bound to the TATA element and its surrounding nucleotides can produce only a low rate of transcription initiation. This rate is enhanced during development or in response to extracellular signals by the binding of transcription factors known as activators to promoter binding sites upstream of the TATA box (Blau *et al.,* 1996, Gill, 2001). In binding to promoter DNA activator proteins increase the rate of transcription initiation by enhancing the recruitment, binding, and stability of general transcription factors involved in the assembly of the pre-initiation complex. Experiments carried out using protein affinity resins have identified several activator proteins that target components of the PIC such as TFIIA, TFIIB, TFIIE, TFIIF, and TFIIH (Berk, 1999, Gill, 2001). The binding sites for several of these activator proteins are found in a wide variety of genes and once bound with the appropriate factor result in increased transcription in all tissues. The SP1 binding site is an example of this and is bound by the cellular activator protein SP1 (Blau *et al.,* 1996).

The HIV-2 LTR contains three consensus SP1 binding sites within the core region of the U3 domain, immediately 5' to the TATA box (Arya & Gallo, 1988, Tong-Starksen *et al.,* 1990) (Figure 1.6B and C, pg 71). Mutagenesis studies have indicated that the three G-C rich sequences extending from -49 to –83 relative to the start site of transcription are functionally important in HIV-2 directed basal and Tat *trans*activated gene expression (Arya & Gallo, 1988, Tong-Starksen *et al.,* 1990, Pagtakhan & Tong-Starksen, 1997). Analyses of the HIV-1 LTR have also revealed the presence of three SP1 binding sites that have functional similarity to those described in the HIV-2 LTR (Nabel & Baltimore, 1987, Harrich *et al.,* 1989). Sequential mutation of the SP-1 sites within the HIV-1 LTR has been shown to result in a progressive decline in basal and Tat-induced gene expression and production of full-length RNA transcripts (Ross *et al.,* 1991, Rittner *et al.,* 1995). Constructs containing mutations in all three HIV-2 SP1 binding sites exhibit severe decreases in HIV-2 gene expression and transcription initiation rates (Pagtakhan & Tong-Starksen, 1997). Moreover, HIV-1 proviral constructs displaying this phenotype are unable to generate progeny virions in some cell types (Harrich *et al.,* 1990, Parrot *et al.,* 1991, Ross *et al.,* 1991, Rittner *et al.,* 1995).

In contrast, reports of the enhanced activity of naturally occurring HIV-1 promoters possessing additional SP-1 sites underline the significant influence that these sites have upon the level of HIV gene expression. Two such reports have demonstrated that natural HIV LTR variants containing four (Koken *et al.,* 1992) and five SP-1 sites (Rousseau *et al.,* 1997) respectively were able to outgrow isogenic constructs

containing three SP-1 binding sites. More recently, Berkhout *et al* described the duplication of the complete SP-1 region through prolonged culturing of an attenuated HIV-1 subtype B virus (Berkhout *et al.*, 1999). This resulted in a stronger LTR promoter with six SP-1 sites, and yielded a fitter virus.

Protein-protein interaction studies have revealed that the SP-1 protein mediates its effect on HIV gene expression by interacting with a subunit of the general transcription factor TFIID, the central member in the pre-initiation complex (Rittner *et al.*, 1995, Blau *et al.*, 1996). By specifically binding to TBP-associated factor $_{II}$110 ($TAF_{II}110$) via its glutamine-rich activation domain, the SP-1 protein is thought to enhance the recruitment, stability, and anchoring of TFIID to the TATA element of the HIV promoter (Hoey *et al.*, 1993, Weinzierl *et al.*, 1993, Rittner *et al.*, 1995). Consistent with a role for the SP-1 protein in the recruitment and stabilisation of TATA binding cellular transcription factors, several studies have demonstrated that the spacing between the SP-1 sites and the TATA box in the HIV-1 and HIV-2 LTRs (Huang & Jeang, 1993, Rittner *et al.*, 1995, Pagtakhan & Tong-Starksen, 1997) is crucial for efficient viral gene expression. Any increases in the distance between the TATA and SP1 sites results in significant decreases in viral gene expression.

In addition to its role in transcription initiation, a number of studies have shown that the SP-1 protein exerts positive effects upon Tat-induced gene expression (Rittner *et al.*, 1995, Blau *et al.*, 1996, Pagtakhan & Tong-Starksen, 1997). Although suggestions of direct interaction between the two proteins have been made, a study by Pagtakhan *et al* has revealed that this is not the case with HIV-2 (Pagtakhan & Tong-Starksen, 1997). While *trans*activation of the HIV-2 LTR by Tat requires *cis*-acting binding sites for SP1, mutagenesis of the activation domains of both proteins has demonstrated that the requirement is not dependent on direct SP1-Tat interaction (Pagtakhan & Tong-Starksen, 1997). The transcriptional synergy observed between the two proteins is thought to be mediated indirectly through another member of the PIC, leading to an overall enhancement in levels of recruitment, stability and processivity of the PIC.

1.10.3.2 The enhancer region of U3

1.10.3.2a.The NF-κB site

The most well studied transcription factor that binds to the sequences within the enhancer region of the HIV-2 U3 domain is NF-κB. Unlike the HIV-1 LTR that possess two functional NF-κB binding elements (Gaynor, 1992), both the HIV-2 and SIV LTRs contain only one 10-bp consensus NF-κB sequence (GGGACTTTCCC), located between nucleotides -110 to –100 (Arya & Gallo, 1988, Renjifo *et al.,* 1990, Tong-Starksen *et al.,* 1990) (Figure 1.6B and C, pg 71).

NF-κB is ubiquitously expressed, but its function is negatively regulated by an inhibitory protein termed IκB (Baeuerle & Baltimore, 1988). In binding to NF-κB, IκB masks a nuclear localisation sequence required for the translocation of the NF-κB protein from the cytoplasm to the nucleus of the cell. Phosphorylation of IκB by protein kinase C (PKC) results in its dissociation from the NF-κB protein leaving it free to translocate into the nucleus of the cell and bind to its promoter recognition sequence (Baeuerle & Baltimore, 1988, Nolan *et al.,* 1991). A number of cell stimuli induce the degradation of IκB including several cytokines (TNF-α and IL-1), T-cell activation signals, and physical/chemical stress (phorbol 12-myristate 13-acetate (PMA), phytohemagglutinin (PHA), and UV etc (Beg & Baldwin, 1993).

Both HIV-1 and HIV-2 transcription is dependent on the integrity of the NF-κB sites within their LTRs. Mutation of the NF-κB motifs in HIV-1 (Kawakami *et al.,* 1988, Ross *et al.,* 1991, Berkhout & Jeang, 1992, West *et al.,* 2001) and HIV-2 (Tong-Starksen *et al.,* 1990, Markovitz *et al.,* 1992, Hannibal *et al.,* 1993, Hilfinger *et al* 1993, Hannibal *et al.,* 1994, Bassuk *et al.*, 1997) LTR CAT constructs has been shown to result in marked decreases in basal and Tat-induced levels of gene expression following transfection.

Because the NF-κB protein plays such a significant role in promoter function, several groups predicted that the HIV-1 promoter, containing two NF-κB sites, would be more active than the HIV-2 promoter containing only one site. It was suggested that this might provide an explanation for the reduced pathogenicity of HIV-2 compared to HIV-1. Comparisons of the transcription levels directed by LTR constructs within T cells activated by pharmacologic agents such as PMA and PHA, revealed little difference between the overall activity of the LTRs from HIV-1 and HIV-2 (Markovitz *et al.,* 1990, Tong-Starksen *et al.,* 1990, Hannibal *et al.,* 1993). Differences were noted, however, when physiologic agents

such as TNF-α, anti-CD3 antibodies, and CD3-specific antigens were used to stimulate T cell activation (Hannibal *et al.,* 1993, Hilfinger *et al.,* 1993). When transfected in the presence of TNF-α, HIV-1 constructs directed a significantly higher transcriptional response to T cell activation compared to the response directed by the HIV-2 LTR. In contrast, transfection studies revealed that when T cells were stimulated by incubation with anti-CD3 antibodies or by antigen specific CD3 receptor recognition, HIV-2 transcription was activated nearly as well as with PMA and PHA, while HIV-1 transcription was induced significantly less than it was by PMA and PHA (Hannibal *et al.,* 1993, Hilfinger *et al.,* 1993, Hannibal *et al.,* 1994). The differential effects of T cell receptor (TCR)-CD3 complex and TNF-α receptor signalling on the HIV LTRs correlated with results in whole-virus replication studies (Markovitz *et al.,* 1990). These studies gave the first indication that the transcriptional regulation of the HIV-2 promoter was different to that of the HIV-1 promoter. It appears that the HIV-1 LTR is stimulated more effectively by TNF-α, a cytokine that acts primarily through NF-κB elements to activate virus replication, in any T cell bearing its receptor. In contrast, the HIV-2 LTR appears to be more effectively stimulated by the calcium-mediated signal transduction pathway activated by soluble anti-CD3 antibodies and TCR mediated antigen recognition (Hannibal *et al.,* 1993, Hilfinger *et al.,* 1993, Hannibal *et al.,* 1994). This method of activation, however, is clonally specific and thus a limited event. It has been suggested that the difference in regulation of the two promoters, rather than a difference in overall activity, may provide a mechanism by which the difference in pathogenicity of the two viruses is determined. However, further evidence outlining the mechanisms by which this may occur is required.

In addition to its response to cellular activation, the NF-κB protein is involved in synergistic interactions with two other activator proteins that play important roles in HIV gene expression, namely SP1 and Tat (Ross *et al.,* 1991, Liu *et al.,* 1992, Perkins *et al.,* 1993, Moses *et al.,* 1994). A study by Ross *et al* has shown that the arrangement and spacing of adjacent NF-κB and SP1 binding sites in the native configuration of the HIV-1 LTR contribute significantly to virus replication (Ross *et al.,* 1991). It has been suggested that SP1 and NF-κB direct a high level of specificity to HIV gene activation by functioning as stereospecific architectural components of enhancer complexes. In this role, the interactions between proteins such as NF-κB, SP1, and Tat enhance transcription factor recruitment and/or binding stability of the already bound transcription factors within the pre-initiation complexes. Thus changes in the relative positions or orientations of the protein-binding sites within the HIV LTR leads to the inactivation of the promoter.

1.10.3.3 The modulatory region of U3

The modulatory region of the U3 domain has been mapped between nucleotides -110 and -180 and contains *cis*-acting transcriptional regulatory elements that, in conjunction with the NF-κB element, modulate the HIV-2 transcriptional response to cellular activation (Markovitz *et al.*, 1990, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995) (Figure 1.6B and C, pg 71). While investigating HIV-2 NF-κB site function in response to T-cell stimulation, the discovery was made that, in contrast to HIV-1, several upstream transcription factor binding sites in addition to NF-κB were required to activate HIV-2 transcription in T-cells (Markovitz *et al.*, 1990, Markovitz *et al.*, 1992). Earlier studies had revealed that mutation of the NF-κB elements within the HIV-1 LTR resulted in the complete abrogation of the transcriptional response to T cell activation (Nabel & Baltimore, 1987). In contrast, the transcriptional response to T cell activation directed by the HIV-2 LTR was merely reduced following mutation of the HIV-2 NF-κB element (Markovitz *et al.*, 1990). Complete inhibition of the transcriptional response was only achieved following the mutation of the sequences upstream of the HIV-2 NF-κB element. These studies demonstrated that the collection of transcription factor binding sites controlling the transcriptional regulation of the HIV-2 LTR were distinct from those found within the HIV-1 LTR, and indicated that transcriptional regulation of the HIV-2 LTR appeared to be a much more complex process than for HIV-1.

At present four upstream modulatory elements have been identified within the HIV-2 LTR, and extensive data have been generated concerning the sequence and function of each of them (Markovitz *et al.*, 1990, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995). Significantly, these elements are distinct from any found within the modulatory region of the HIV-1 LTR, but are conserved among HIV-2 isolates. Moreover, comparative analysis of the HIV-2 and SIV modulatory promoter regions has revealed 75% homology, with SIV LTR sequences containing at least two of the four *cis*-acting inducible transcription factor binding elements found within the HIV-2 LTR (Renjifo *et al.*, 1990, Pöhlman *et al.*, 1998). Most interestingly, detailed analyses of the sites within the modulatory region have shown that they act synergistically to control both basal and activated levels of transcription (see section 1.10.3iiic).

1.10.3.3a. The PuB1, PuB2, and peri-ets (pets) elements

The first HIV-2 modulatory site to be identified was termed purine rich box 1 (PuB1) located between nucleotides -174 and -159 relative to the start site of transcription (Figure 1.6C, pg 71). In 1990

Markovitz and co-workers demonstrated that mutation of the sequences between nucleotides –174 to –159, containing the putative PuB1 site, completely abrogated the transcriptional response to PMA, directed by a ΔNFκB-HIV-2 LTR (Markovitz *et al.*, 1990). This effect was also seen when other pharmacologic stimulants such as phytohemagglutinin (PHA) were used to induce T-cell activation.

Following the discovery of the PuB1 element, several internal LTR deletion and substitution mutants were generated to examine the importance of the spacing between the PuB1 and NF-κB sites. In characterising these mutations further transcriptional decreases were noted and lead to the discovery of two additional *cis*-acting elements required for the regulation of the HIV-2 LTR in response to T-cell stimulation (Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994). The first *cis*-acting element was a purine-rich sequence conserved amongst HIV-2 and SIV isolates and located between nucleotides –142 and –137 relative to the start of transcription (see Figure 1.6B and C, pg 71). This element shared significant homology with the PuB1 site and was subsequently termed purine-rich box 2 (PuB2) (Markovitz *et al.*, 1992). The next element to be identified was a TG-rich element (TTGGTCAGGG) located immediately upstream of the PuB2 site termed the proximal-*ets* (pets) site (Markovitz *et al.*, 1992) (see Figure1.6 B. and C, pg 71).

The PuB1 and PuB2 elements within the HIV-2 promoter have been shown to bind proteins belonging to the Ets family of proto-oncogenes (Leiden *et al.*, 1992, Markovitz *et al.*, 1992). Ets binding motifs usually provide the recognition sequence for the ets binding factor Ets-1, and are commonly found in other type C mammalian retroviruses, excluding HIV-1 (Markovitz *et al.*, 1992). However, Ets-1 was found not to recognise either PuB1 or PuB2. Instead recombinant studies revealed that both purine rich elements bound E74-like factor-1 (Elf-1) a member of the Ets proto-oncogene family which is related to the *Drosophila* E74 transcription factor (Bhat *et al.*, 1989, Ben-David *et al.*, 1991, Thompson *et al.*, 1991, Leiden *et al.*, 1992, Markovitz *et al.*, 1992, Thompson *et al.*, 1992, Wang *et al.*, 1992). In particular, Elf-1 binds to the core pentanucleotide motif (AGGAA), located within each of the purine-rich sequences of PuB1 and PuB2 (see Figure 1.6C, pg 71).

DNase footprinting analysis and electrophoretic mobility shift assays (EMSAs) performed with the pets site have found that this site is constitutively bound by an autoantigen termed DEK, a protein usually involved in chromosomal translocation in acute myelogenous leukaemia (Fu & Markovitz, 1996, Fu *et al.*, 1997). More recent analyses have indicated that intracellular levels of DEK can be modulated in response to cell stimulation and differentiation by phorbol esters such as 12-*O*-tetradecanoylphorbol-13-

acetate (TPA) (Faulkner *et al.*, 2001). TPA acts to stimulate signalling cascades such as the protein kinase C (PKC) pathway that are usually triggered in response to T-cell receptor activation by antibody or antigen. The study by Faulkner *et al* demonstrated that in response to TPA, a protein phosphatase PP2A triggers the dephosphorylation of DEK through an interaction with the PKC pathway resulting in its exchange for another as yet unidentified pets binding factor (Faulkner *et al.*, 2001). This exchange is accompanied by an increase in the level of transcription directed by the HIV-2 LTR. Interestingly, okadaic acid (OKA) a well-known inhibitor of PP2A has been shown to activate HIV-1 transcription, leading to the assumption that PP2A would suppress HIV-1 transcription. The finding that PP2A appears to boost transcription regulated by the HIV-2 LTR could represent another difference of transcriptional regulation between the two viruses (Faulkner *et al.*, 2001).

Mutation of the PuB1, PuB2, NF-κB or pets site has been shown to reduce markedly basal transcription and the induction of transcription from the HIV-2 LTR following T-cell stimulation (Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hannibal *et al.*, 1994). Alteration of these *cis*-acting elements leads to correctly initiated, although markedly diminished; levels of RNA indicating that the sites act to determine the rate of transcription initiation. In addition to their role in the transcriptional response to T cell stimulation, transfection analysis in monocytes at various stages of differentiation has revealed that the PuB1, PuB2 and pets sites are required along with the NF-κB element, to direct basal and activated levels of gene expression in these cell lines (Hilfinger *et al.*, 1993). This is again in contrast to HIV-1 regulation in which developmentally controlled expression from the LTR in monocytes, as in T cells, is mediated largely through the NF-κB sites alone (Griffin *et al.*, 1989).

1.10.3.3b. The Peri-κB element

While characterising modulatory site function in immature and mature monocytes a fourth *cis*-acting element was identified and shown to be required for the function of the HIV-2 LTR following monocytic stimulation (Clark *et al.*, 1995). The peri-κB element, named for its close proximity to the NF-κB site, is located between nucleotides -127 and -109, and is highly conserved among isolates of HIV-2 and closely related strains of SIV (see Figure 1.6B and C, pg 71). Transfection studies have demonstrated that this site displays a unique property not identified in any of the previously described *cis*-acting elements, peri-κB functions to influence HIV-2 LTR activity in a cell-type specific manner. LTR constructs containing changes to the nucleotides encompassing the peri-κB sequence direct reduced levels of basal and

activated transcription in monocytic cell lines, but direct levels of gene expression indistinguishable from wild type promoter constructs when transfected into T cell lines (Clark *et al.*, 1995).

Analysis of the peri-κB site has revealed that an as yet unidentified nuclear factor is bound to the site in both peripheral blood monocytes and T cells. It has been proposed that the cell specificity of the peri-κB site is determined either by a different protein binding to this site in monocytes versus T cells or by differential modification of the binding protein in monocytes and T cells (Clark *et al.*, 1995).

1.10.3.3c. Cis-acting binding factors and transcriptional synergy

Following the identification of the inducible *cis*-acting elements upstream of the HIV-2 NF-κB site much work was put into characterising the mechanism by which these sites exerted their functional influence upon the rate of transcription directed by the HIV-2 LTR. When inserted upstream of the simian virus 40 promoter, multiple copies of the PuB1 site were found to be insufficient to confer PMA or PHA responsiveness within T-cells. This indicated that PuB1 required the presence of the other *cis*-acting enhancer elements to provide PMA responsiveness to the HIV-2 LTR. Further mutation analyses confirmed that the functional effects exerted by the *cis*-acting elements within the modulatory region of the HIV-2 LTR, at the basal and activated levels, required the collective presence of each of the four elements (Markovitz *et al.*, 1992, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994). It has been suggested that the transcriptional synergy displayed by the modulatory elements within the HIV-2 LTR is mediated by specific interactions between the transcription factors that recognise and bind to them.

Several experimental systems have shown that members of the Ets family of proto-oncogenes readily interact with other transcription factors involved in promoter activation, and in doing so enhance either general transcription factor recruitment or stabilize their binding (Wang *et al.*, 1994, John *et al.*, 1995, Sun *et al.*, 1995). Moreover, synergistic regulation by two of the transcription factors that bind to the modulatory elements in the HIV-2 LTR has been shown to control, at least in part, transcription directed by the HTLV-1 enhancer (Clark *et al.*, 1993). The study by Clark *et al* has demonstrated functionally significant interactions between the Ets-related PuB binding factor, and the pets binding factor DEK (Clark *et al.*, 1993). It has been postulated that the transcriptional synergism exerted by the *cis*-acting inducible elements within the HIV-2 LTR is determined by the interactions between, Elf-1, bound to PuB1 and PuB2, and the pets binding factor and NF-κB proteins bound to the pets and NF-κB sites

respectively (see Figure 1.7, pg 84). In the case of monocytes, the peri-κB binding factor also appears to be involved in, and required for the synergistic induction of transcription from the HIV-2 LTR.

Thus, the transcription factors which bind to the *cis*-acting modulatory elements within the HIV-2 LTR function as highly stereospecific architectural components of 3-D nucleoprotein transcriptional complexes. The influence that they exert upon the level of HIV-2 gene expression is dependent on the relative positions of their binding sites within the promoter and upon their orientation to the proteins with which they interact. In binding to the *cis*-acting sites within the modulatory region of the LTR the transcription factors are thought to enhance the recruitment of members of the PIC and/or additional cofactors to the promoter, and to increase the stability of transcription factors that are already bound to the HIV-2 LTR (see Figure 1.7, pg 84).

Binding of the PuB1 and PuB2 binding factor, Elf-1, is constitutive (Markovitz *et al.*, 1992, Leiden *et al.*, 1992, Fu & Markovitz, 1996, Fu *et al.*, 1997), and does not appear to change following cellular stimulation, therefore posttranslational modification of this DNA-binding protein via phosphorylation/dephosphorylation is likely required for translational regulation (Fu & Markovitz, 1996, Fu *et al.*, 1997). Such modifications have already been shown to play an important role in the regulation of HIV-2 transcription by DEK and the second unidentified pets binding factor (Faulkner *et al.*, 2001). The posttranslational changes are likely to affect the stereospecific interaction between the proteins bound to the HIV-2 LTR in a way that promotes an increase in the rate of transcription initiation via transcription or cofactor recruitment and stability. This results in an overall increase in HIV-2 gene expression (Hilfinger *et al.*, 1993, Clark *et al.*, 1995, Fu & Markovitz, 1996, Fu *et al.*, 1997, Faulkner *et al.*, 2001).

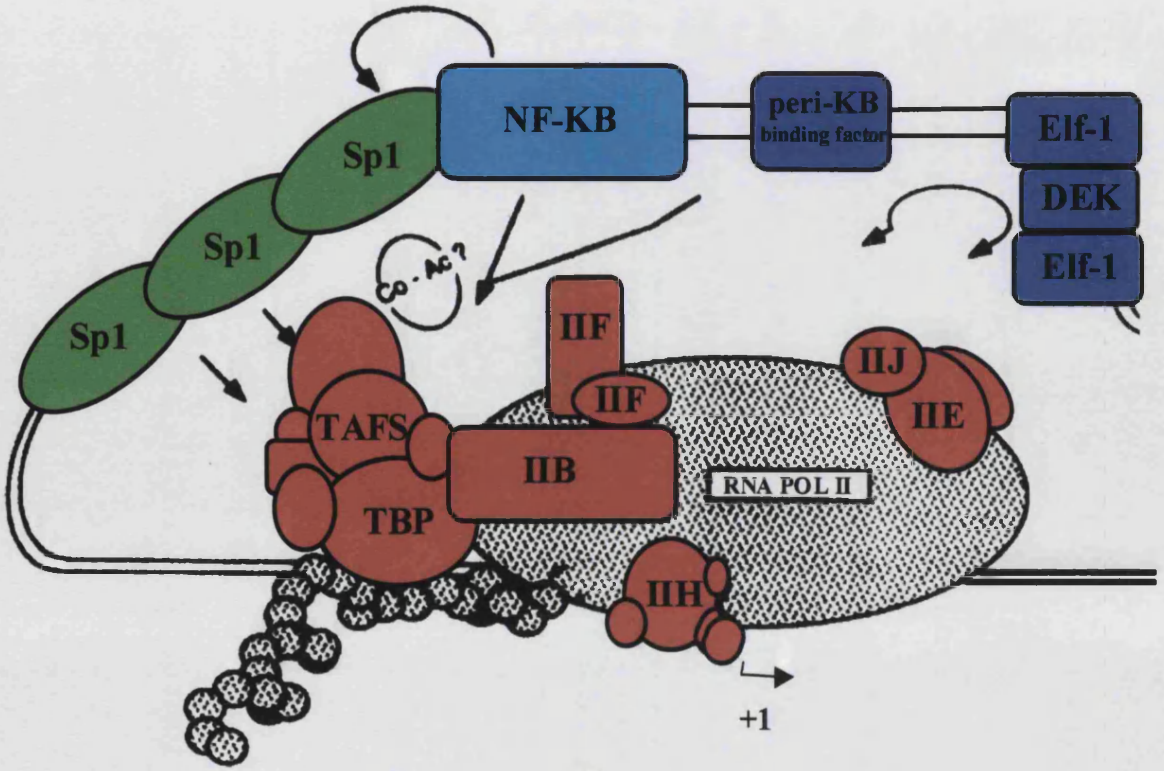


Figure 1.7 Model for transcriptional activation of the HIV-2 LTR by upstream factors. Initiation of HIV-2 transcription depends upon the binding of general transcription factors TBP and TAFs (see section 1.10.1) to the promoter. Transcriptional activation of the HIV-2 LTR also requires the transcription factors SP1, NF-κB, Elf-1, DEK, and peri- κB binding proteins to bind to their respective recognition elements within the enhancer and modulatory regions of the promoter. In doing so the upstream factors and SP1 help to recruit general transcription factors (TFIIA-J and RNAPII) to the LTR and stabilise the preinitiation complex (PIC). Activation of transcription by the Elf-1, DEK, and peri- κB binding proteins has been shown to be synergistic, thus in addition to the interactions these proteins share with general transcription factors they also require interaction with each other. In these positions, the upstream factors that bind to the HIV-2 LTR act as stereospecific architectural components of the transcriptional complex. (Adapted from Jones & Peterlin, 1994).

1.10.4 The structure and function of the binding elements within the R domain of the HIV-2 LTR

1.10.4.1 The trans-activation response (TAR) RNA element

The R region of the HIV-2 LTR extends from nucleotide +1 to + 173 relative to the start of transcription (Figure 1.6A, pg 71). Chemical and enzymatic analyses of this region have demonstrated that it encodes the recognition site for the viral *trans*activator protein Tat and a number of associated cellular proteins (Guyader *et al* 1987, Arya & Gallo, 1988, Fenrick *et al.,* 1989, Chang & Jeang, 1992, Elangovan *et al.,* 1992, Rhim & Rice, 1993). In the presence of Tat the level of gene expression directed by the HIV-2 LTR is markedly increased in a process called *trans*activation (see section 1.11.1). The Tat responsive sequence responsible for mediating this effect has been mapped between nucleotides +1 and +123 of R relative to the start of transcription (Emerman *et al.,* 1987, Guyader *et al.,* 1987); this sequence encodes a 3-D RNA structure termed the HIV-2 *transa*ctivation *r*esponse element or TAR. The HIV-2 TAR element folds into a structure containing three stem-loop hairpins and can be found at the 5' end of all HIV-2 nascent transcripts (Figure 1.8, pg 85). In contrast, the HIV-1 TAR (TAR1) element is smaller than that of HIV-2 (TAR2) at only 45 nucleotides in length, and configures into a less complex single stem-loop structure (Figure 1.8, pg 85).

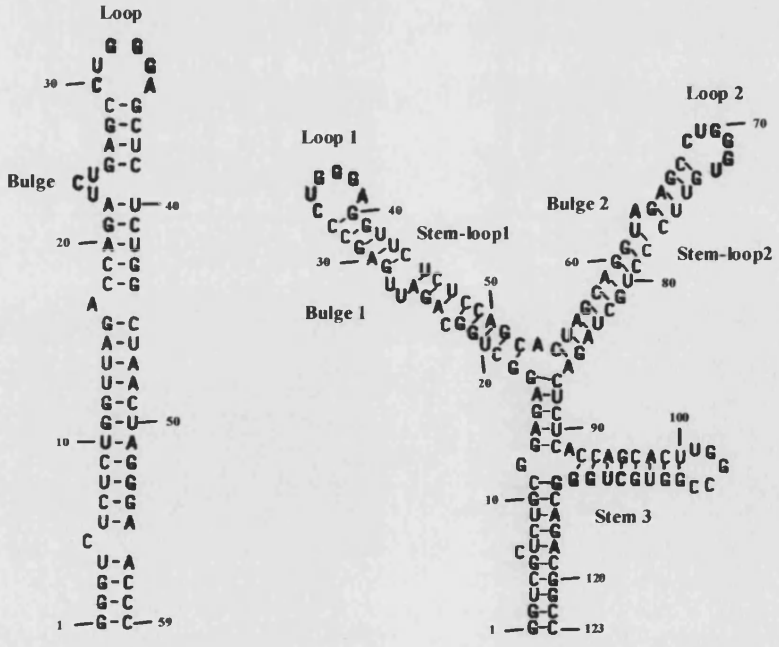


Figure 1.8 Comparison of the primary and predicted secondary structures of HIV-1 TAR and HIV-2 TAR (TAR2). The critical functional elements in TAR are the terminal hexanucelotide loops and the 5' proximal bulges. The position of the loop and bulge sequences in both the HIV-1 and HIV-2 TARs are indicated. (adapted from Garcia-Martinez *et al.,* 1995).

Extensive investigations into the function of TAR have revealed that it acts as the binding site for the Tat protein and its associated host cell proteins during the process of transcriptional *trans*activation, properly positioning them for functional interactions with the transcription factors of the pre-initiation complex (PIC) (see section 1.11.1) (Arya & Gallo 1988, Berkhout *et al.,* 1990, Rhim & Rice, 1993).

The three stem-loop structures of TAR2 are formed by the basepairing of the nucleotides between position +1 and +123 relative to the start site of transcription (see Figure 1.8, pg 85). Deletion analyses of the nucleotides within these three structures have revealed that the first stem-loop structure (stem-loop1) is more critical for Tat *trans*activation than the second stem-loop structure (stem-loop2), and that the third stem-loop structure is not required at all (Fenrick *et al.,* 1989, Rhim & Rice, 1993, Browning *et al.,* 1997). Stem-loop2 has been shown to mediate Tat *trans*activation, all be it at a reduced level (30% that of WT), in the presence of an altered stem-loop1 (Fenrick *et al.,* 1989, Berkhout *et al.,* 1990, Garcia-Martinez *et al.,* 1995). However, in the context of a fully functioning stem-loop1, mutation of stem-loop2 has little or no measurable effect on transcription levels in the presence of Tat (Fenrick *et al.,* 1989, Rhim & Rice, 1993, Browning *et al.,* 1997). Fenrick *et al* has suggested that the limited involvement in *trans*activation of stem-loop2 is due to its distance from the cap site, since decreasing the distance by deletion significantly increased *trans*activation of stem-loop2 by Tat2 (Fenrick *et al.,* 1989).

Extensive analyses of the sequence and secondary structure of TAR2 have revealed a number of functionally significant features. Both stem-loop1 and stem-loop2 possess highly conserved dinucleotide bulges within the promoter-proximal arm of the stem regions, in addition to conserved unpaired hexanucleotide loop motifs at the apical tips of the hairpin structures (Fenrick *et al.,* 1989, Rhim & Rice, 1994a). These features are similar in position, sequence, and structure to the trinucleotide bulge and hexanucleotide loop motifs found within the single hairpin structure of TAR1 (Rhim & Rice, 1994a). As with HIV-1 TAR, the dinucleotide bulges of the hairpin structures in TAR2 have been shown to provide the binding motifs for the HIV-2 Tat protein (Rhim & Rice, 1993, Rhim & Rice, 1994a). More recent analysis has suggested that the nucleotide basepairs immediately above and below the bulge motif are also involved in the Tat2-TAR2 binding interaction (Berkout *et al,* 1990, Chang & Jeang, 1992, Churcher *et al.,* 1993, Brodsky *et al.,* 1998). Studies of the solution structure of TAR1 indicate that the bulge motif of TAR induces a bend in the RNA helix that distorts the local structure and widens the major groove in the double stranded RNA stem to expose hydrogen-bonding contacts that are important for stabilising the binding of Tat (Weeks *et al.,* 1990, Colvin & Garcia, 1992, Delling *et al.,* 1992).

Protein binding studies have demonstrated that deletion of the dinucleotide bulge motifs within stem-loop1 and 2 markedly reduces Tat2 binding *in vitro* and largely abolishes the *trans*activation of transcription via the TAR2 element *in vivo* (Rhim & Rice, 1994a, Garcia-Martinez *et al.*, 1995). Removal of the bulge within stem-loop1 only, reduces Tat *trans*activation *in vivo* by approximately 70%, while deletion of the bulge in stem-loop2 has no measurable effect on transcription levels. Nonetheless, removal of either dinucleotide bulge has been shown to simplify the *in vitro* Tat2-TAR2 binding pattern, suggesting that other interactions in addition to those directed by the dinucleotide bulge of stem-loop1 are involved in the recruitment of Tat2 to TAR2 (Dingwall *et al.*, 1990, Rhim & Rice, 1994a, Garcia-Martinez *et al.*, 1995, Browning *et al.*, 1997).

While mutagenesis of the TAR hexanucleotide loop sequences CUGGGA and CUGGGU was not found to result in a loss in Tat binding, preliminary functional analyses of these mutants revealed that Tat *trans*activation of transcription was completely inhibited. This finding indicated that a loop-binding factor in addition to the Tat2 protein was required to induce Tat *trans*activation of transcription. After an extensive search, cyclin-T1 (CycT1), a subunit of the positive transcription elongation factor complex (P-TEFb), has been identified as the cellular cofactor that recognises the loop sequences of TAR *in vivo* (Bieniasz *et al.*, 1998, Fujinaga *et al.*, 1998, Garber *et al.*, 1998a, Garber *et al.*, 1998b, Wei *et al.*, 1998). In particular, studies using the HIV-1 TAR loop sequence as a model have demonstrated that the first and third guanidine residues within the hexanucelotide loop sequence are critical for the recruitment of CycT1 to TAR (Richter *et al.*, 2002a, Richter *et al.*, 2002b). The identification of the TAR loop-binding cofactor has lead to a greater understanding of the mechanism by which Tat causes an increase in the level of HIV gene expression (see section 1.11.1).

In addition to the functional motifs of TAR2, proper folding of the stem-loop structures is required. A study characterising the effects of disrupting the basepairing in the upper and lower stems of stem-loops 1 and 2 showed marked reductions in Tat-induced transcription and simplified binding patterns (Browning *et al.*, 1997). This study has indicated that more yet unidentified cellular factors may bind to the TAR region of HIV-2, some of which may be involved in mediating the Tat response *in vivo*. Further analyses are required to fully characterise the HIV-2 TAR region and the factors for which it is a target.

The TAR element has been found to play several roles in other steps of the viral life cycle (Klaver & Berkhout, 1994). Specifically, both the 5' and 3' TAR structures have been shown to contribute to optimal packaging of the viral genome into virion particles (McBride & Panganiban, 1997, McBride *et*

al., 1997, Das *et al.,* 1998, Harrich *et al.,* 2000). In addition the TAR hairpins have been reported to play a role in reverse transcription, however these additional TAR functions have yet to be fully explored (Harrich *et al.,* 1996, Das *et al.,* 1998).

1.10.5 The structure and function of the binding elements within the U5 domain of the HIV-2 LTR

Unlike the U3 and R domains of the HIV-2 LTR, deletion analyses performed on the U5 domain have failed to reveal the presence of functionally significant transcription factor binding sites (Arya, 1991). A study by Arya reported that the sequences within U5 could downmodulate HIV-2 LTR-directed gene expression, however this effect appeared to be dependent upon the length of U5 rather than the identity of the nucleotides within the sequence (Arya, 1991). Furthermore, the repressive effect of the downstream sequence of U5 appeared to be cell-type dependent. Finally, the downmodulation was not observed at the Tat-induced or T-cell activated level of gene expression, indicating perhaps only a minor role for these sequences in the regulation of HIV-2 gene expression (Arya, 1991).

1.11 The HIV-2 Tat protein

Of the known lentiviruses, HIV-1, HIV-2, SIV, and EIAV encode Tat proteins that markedly increase levels of transcription through leader TAR RNA structures (Jones & Peterlin, 1994). The Tat protein of HIV-2 is a potent *trans*activator of transcription from the HIV-2 LTR and has been shown to be essential for efficient viral replication. Full length HIV-2 Tat (ROD isolate) is expressed early in infection and is encoded by two separate exons, producing a protein 130 amino acid residues in length (Arya *et al.*, 1987, Emerman *et al.*, 1987).

Genetic and biochemical experiments have established that, like many transcriptional activators, Tat2 has a modular structure, with an RNA binding domain and a strong transcriptional activation domain (Chang & Jeang, 1992, Arya, 1993, Rhim & Rice, 1994b, Pagtakhan & Tong-Starksen, 1995) (see Figure 1.9, pg 89). The RNA binding domain consists of a stretch of positively charged basic amino acid residues that direct Tat to the HIV-2 TAR RNA element where it binds in a 1:1 complex (Elangovan *et al.*, 1992, Jones & Peterlin, 1994). Positively charged amino acid residues encompassed by the arginine-rich basic domain interact with the negatively charged phosphate backbone of the TAR RNA determining a high Tat:TAR binding affinity (Frankel, 1992, Gait & Karn, 1993). The basic region within the RNA binding domain of Tat also contains a nuclear localisation signal, which functions to target the protein to the cell nucleus following expression.

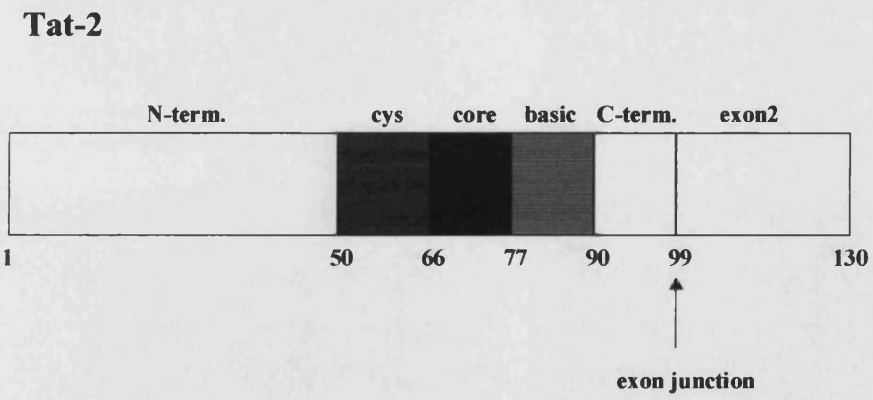


Figure 1.9 Schematic representation of the two exon Tat-2 (HIV-2 $pROD_{10}$) protein. The junction between the exon 1 and exon 2 is indicated. The TAR binding domain is located within the basic region in exon 1. This region also serves as a nuclear localisation signal. The independent activation domain of Tat2 includes the amino-terminal, cysteine-rich, and core regions of the protein (Rhim & Rice, 1994b, Pagtakhan & Tong-Starksen, 1995).

The activation domain of Tat2 can be subdivided into three regions: the non-conserved amino-terminal domain, the cysteine-rich domain, and the hydrophobic core domain (Arya, 1993, Rhim & Rice, 1994b, Pagtakhan & Tong-Starksen, 1995) (see Figure 1.9, pg 89). The amino-terminal domain contains several non-conserved acidic amino acids that are predicted to form an amphipathic α-helix. This region also contains a number of proline residues; which are presumed to introduce turns in this structural element. The cysteine-rich domain as its name suggests contains seven highly conserved cysteine residues, of which six are critical for Tat2 function (Sadaie *et al.*, 1988, Arya, 1993, Rhim & Rice, 1994b, Pagtakhan & Tong-Starksen, 1995), these residues have been shown to be involved in intramolecular disulphide bridge formation (Koken *et al.*, 1994). The central core region is the most conserved domain of Tat2 and contains residues which have been shown to enhance the affinity and specificity of the Tat:TAR interaction both *in vitro* and *in vivo*.

The second exon of Tat2 encodes a conserved carboxy-terminal domain that has been shown to contribute to optimal *trans*activation of the HIV-2 LTR (Arya, 1993, Tong-Starksen *et al.*, 1993). A study by Rhim & Rice demonstrated that this exon contributed to *trans*activation of the HIV-2 LTR by increasing the binding affinity to HIV-2 TAR RNA (Rhim & Rice, 1994b).

1.11.1. Mechanism of Tat action

The identification of the stage of transcription at which Tat exerts its function, and the mechanism by which it markedly increases the level of gene expression has been widely investigated. Although Tat has been shown to have modest effects on the initiation of RNA synthesis, it has now been established that Tat functions to enhance the efficiency of transcriptional elongation (Marciniak & Sharp, 1991, Berkhout & Jeang, 1992, Kato *et al.*, 1992, Lapsia *et al.*, 1993, García-Martínez *et al.*, 1995, Rittner *et al.*, 1995, Blau *et al.*, 1996, Fujinaga *et al.*, 1998).

In the absence of Tat, the majority of correctly initiated RNAPII transcription complexes stall or terminate transcription prematurely, 20 - 40 nucleotides downstream of the HIV LTR (Kao *et al.*, 1987, Kessler & Mathews, 1992) (Figure 1.10A, pg 94). This gives rise to an abundance of short or non-processive transcripts but relatively few full-length transcripts. In contrast to activator proteins such as SP1 which enhance the rate of transcription initiation and skew polymerase distribution with a sharp drop in density immediately downstream from the 5' end of the template, Tat acts to enhance the processivity of paused RNAPII complexes and gives rise to a uniform distribution of polymerases along the gene, resulting in a large increase in the number of full-length transcripts and relatively few truncated

transcripts (Feinberg *et al.*, 1991, García-Martínez *et al.*, 1995, Rittner *et al.*, 1995, Blau *et al.*, 1996, Bourgeois *et al.*, 2002).

1.11.1.1 Tat interacts with a CTD kinase to phosphorylate the RNAPII CTD

As described in section 1.10.1, studies of eukaryotic transcription have revealed that the release of the stalled RNAPII transcription complexes and their transition into elongation complexes requires the phosphorylation of the C-terminal domain (CTD) of the largest RNAPII subunit. Extensive analyses of the HIV *trans*activation process have revealed that activation of transcription by Tat also involves the CTD of RNAPII (Chun & Jeang, 1996, Okamoto *et al.*, 1996, Parada & Roeder, 1996, Yang *et al.*, 1996).

Initial findings suggested that, once anchored to the promoter by the TAR element, the HIV Tat protein enhanced the processivity of stalled transcription complexes by directly interacting with the CTD of RNAPII. However, in finding that TAR-dependent *trans*activation by Tat was blocked effectively by the kinase inhibitor 5, 6-dichloro-1-β-D-ribofuranosyl-benzimidazole (DRB), several studies indicated that it was more likely that the interaction between Tat and TAR affected a CTD kinase-dependent step in the elongation process (Marciniak & Sharp, 1991).

Following the identification of a Tat-associated kinase (TAK) that bound specifically to the Tat protein *in vitro*, it was postulated that, the activation domain of Tat acted to recruit a kinase to the promoter in binding to TAR, resulting in the phosphorylation of the CTD of RNAPII (Herrmann & Rice, 1993, Herrmann & Rice, 1995). In accordance with this, a study by Yang and co-workers in 1996 demonstrated that the Tat-TAK interaction correlated precisely with the ability of wild type and mutant HIV-1, HIV-2, and EIAV Tat proteins to stimulate transcription *in vivo* (Yang *et al.*, 1996).

The actual identity of the Tat-associated kinase followed the discovery of the human form of the *Drosophila* positive elongation factor b (P-TEFb). Microsequence analyses of human P-TEFb revealed that it was indistinguishable from the Tat-associated kinase complex-TAK (Cujec *et al.*, 1997, Yang *et al.*, 1997, Zhu *et al.*, 1997). Positive elongation factor b (P-TEFb) is composed of at least two subunits: the catalytic subunit cyclin-dependent kinase CDK9 (previously named PITALRE) and the regulatory subunit cyclin T1 (CycT1) (Yang *et al.*, 1997, Zhu *et al.*, 1997, Wei *et al.*, 1998). Following its discovery several groups went on to demonstrated that P-TEFb bound specifically to the activation domain of Tat, and that this interaction was eliminated by mutations that abrogated Tat-induced transcription (Zhu *et al.*, 1997). Furthermore, P-TEFb was found to bind to the Tat proteins of HIV-2 and EIAV which both

appear to function through the same cofactor as HIV-1 Tat (Carroll *et al.*, 1992, Madore & Cullen, 1993). Antibodies directed against a 20-residue peptide from the carboxyl terminus of P-TEFb simultaneously removed TAK activity and rendered HeLa nuclear extracts dependent on P-TEFb for efficient elongation. Most importantly, HNE depleted of P-TEFb were unable to support Tat-induced transcription. Finally, phosphorylation of the RNAPII CTD by P-TEFb was found to be DRB-sensitive.

Several groups have since confirmed the observations made by Zhu *et al*, and have demonstrated that P-TEFb is required for HIV Tat *trans*activation *in vivo* and *in vitro* (Mancebo *et al.*, 1997, Fujinaga *et al.*, 1998, Zhou *et al.*, 1998, Bieniasz *et al.*, 1999a). As in eukaryotic elongation, the CDK9 kinase subunit of P-TEFb acts to hyperphosphorylate the CTD of RNAPII at serine 2 and serine 5 (Zhou *et al.*, 2000), converting the pre-initiation complex into a highly processive elongation complex (Kim *et al.*, 2002) (Figure 1.10B, pg 94). Because P-TEFb is a limiting transcription factor and is not associated with RNAPII initiation or elongation complexes (Marshall & Price 1995, Marshall *et al.*, 1996) Tat is thought to increase the rate of elongation by recruiting P-TEFb to RNAPII through binding to TAR, effectively increasing the local concentration at the promoter (Garcia-Martinez *et al.*, 1997a, Garber *et al.*, 1998a, Bieniasz *et al.*, 1999a, Ping &Rana, 1999). The interactions between Tat, P-TEFb and TAR are thought to induce conformational changes in the CDK9 enzyme that result in its activation (Bieniasz *et al.*, 1998, Chen *et al.*, 1999, Fujinaga *et al.*, 1999, Garber *et al.*, 1998b, Ivanov *et al.*, 1999, Kwak *et al.*, 1999), the subsequent phosphorylation of the CTD of RNAPII by CDK9 (Kim *et al.*, 2002), and the release of the elongating complex from TAR (Keen *et al.*, 1997). In addition, Tat-activated CDK9 has been shown to phosphorylate the positive elongation factor Spt5 in parallel to the hyperphosphorylation of the CTD of RNAPII during HIV-1 transcription (Wu-Baer *et al.*, 1998, Ivanov *et al.*, 2000, Kim & Sharp, 2001, Bourgeois *et al.*, 2002). Once phosphorylated Spt5 acts as a complementary factor for Tat-activated transcription by inhibiting premature RNA release from the numerous pause sites within the HIV genome. The exact mechanism by which CDK9 becomes activated by Tat however, remains unclear (Kim *et al.*, 2002).

Recent studies have demonstrated that the affinity and specificity of the Tat-TAR interaction is governed by the cyclin T1 protein associated with CDK9 within the P-TEFb complex. CDK9-associated cyclin T1 interacts directly with the Tat activation domain when Tat binds to P-TEFb and confers the requirement for the hexanucelotide loop sequences of TAR that mutational analysis has demonstrated are required for Tat-induced activation of transcription. Several studies have demonstrated that cyclin T1 is the TAR RNA loop-binding cofactor for Tat and have shown that, while required for *trans*activation by Tat,

neither CDK9 nor cyclin T1 can bind TAR alone. The Tat-CDK9-cyclin T1 heterodimer has been shown to have a far-higher affinity for TAR than Tat alone (Bieniasz *et al.*, 1998, Garber *et al.*, 1998a, Garber *et al.*, 1998b, Wei *et al.*, 1998), and has lead to the conclusion that Tat complexes with P-TEFb before binding to TAR, leading to CDK9-directed phosphorylation of the CTD of RNAPII and cyclin T1-enhanced Tat-TAR binding avidity and stability (Figure 1.10B, pg 94).

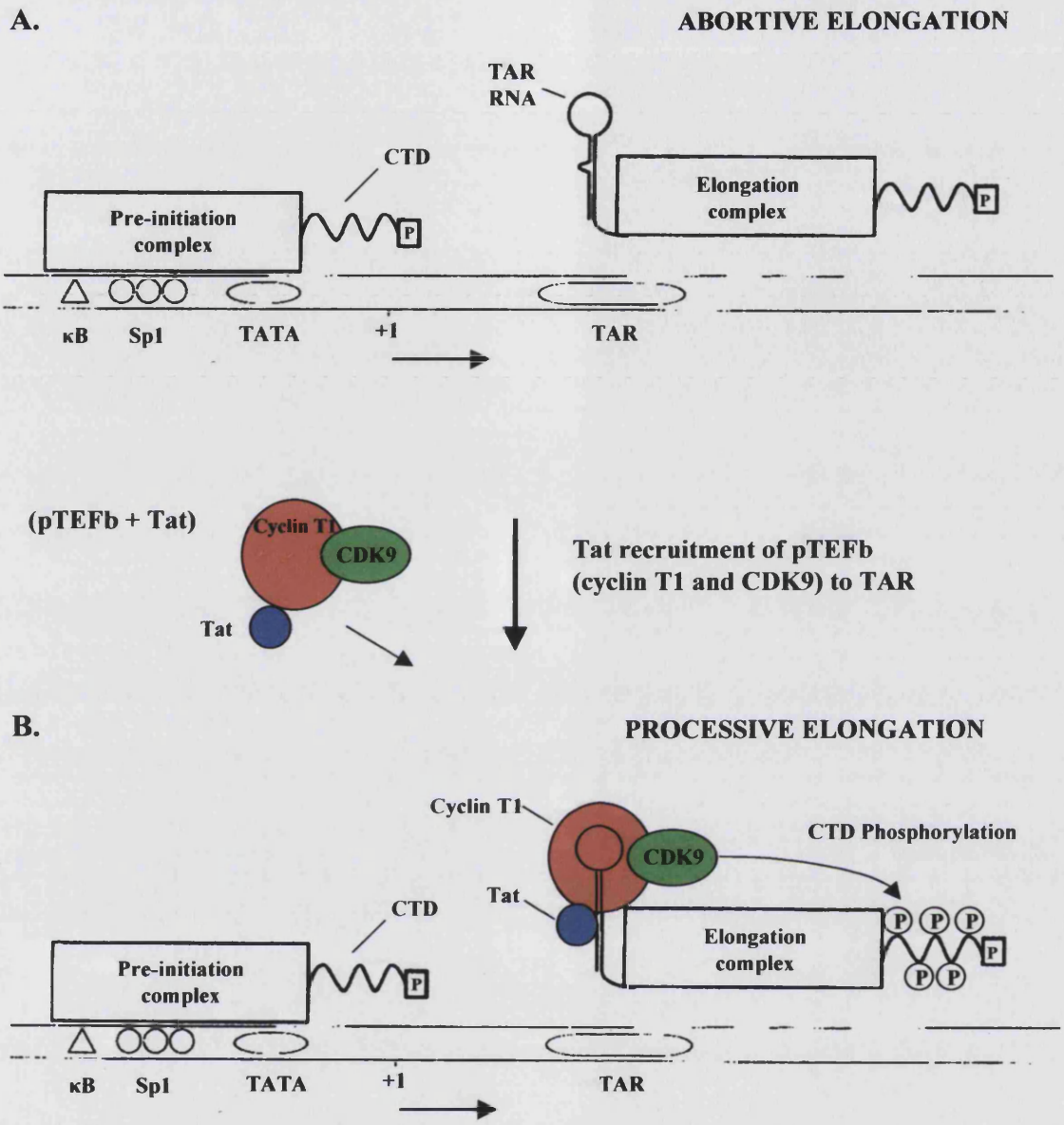


Figure 1.10 Diagram showing the mechanism of *trans*activation by Tat. (A). In the absence of Tat, elongation complexes are not highly processive and pause or abort transcription 20-40 nucleotides downstream from the start site of transcription (Kao *et al.*, 1987, Kessler & Mathews, 1992). (B). It is proposed that Tat enhances the processivity of paused RNAPII complexes by recruiting the human positive elongation factor pTEFb to the TAR element (Garcia-Martinez *et al.*, 1997a, Garber *et al.*, 1998a, Bieniasz *et al.*, 1999a, Ping &Rana, 1999). The cyclin T1 subunit of pTEFb mediates the loop sequence specificity to the Tat-pTEFb complex and in binding to Tat enhances Tats binding affinity for TAR (Bieniasz *et al.*, 1998; Garber *et al.*, 1998a; Garber *et al.*, 1998b; Wei *et al.*, 1998). Once bound to TAR, the cyclin-dependent kinase 9 (CDK9) subunit of pTEFb hyperphosphorylates the CTD of RNAPII, inducing conformational changes to the enzyme that result in the formation of a highly processive elongation complex and an increase in the number of full-length transcripts (Zhou *et al.*, 2000, Kim *et al.*, 2002).

The identification of CycT1 as the loop-binding cofactor for Tat also led to a molecular explanation for several key aspects of Tat function. For example, to function the Tat binding domain has been shown to require the presence of the activation domain; in complexing with P-TEFb the activation domain of Tat recruits the loop-binding cofactor in addition to its *trans*activating kinase subunit, thereby mediating binding specificity as well as functional specificity. In addition, transfected rodent cells will only support low levels of Tat *trans*activation as a result of poor TAR recognition. A recent study by Fujinaga and co-workers comparing the amino acid sequences of cloned human and mouse cyclin T has revealed that a single point mutation at position 261 in mouse cyclin T is responsible for this phenotype (Fujinaga *et al.*, 1999). Introduction of human cyclin T into rodent cells increases levels of Tat *trans*activation, and the reciprocal exchange of the tyrosine at position 261 to the cysteine found in human cyclin T renders mouse cyclin T fully functional for the effects of Tat in cells (Garber *et al.*, 1998b, Fujinaga *et al.*, 1999).

Previously, several groups have provided data demonstrating that the Tat2 protein is only poorly able to activate transcription via the HIV-1 TAR element, (Arya *et al.*, 1987, Emerman *et al.*, 1987, Fenrick *et al.*, 1989, Berkhout *et al.*, 1990) yet the Tat1 protein is able to fully *trans*activate TAR2 in addition to its own TAR1 target despite differences in TAR1 and TAR2 structure. Yeast-two hybrid studies analysing cyclin T interactions with both the Tat1 and Tat2 proteins have revealed that these two proteins differ significantly in terms of their TAR RNA sequence specificity. The human cyclin T1-Tat1 complex binds to both TAR1 and TAR2 with equal efficiency, in contrast the human cyclin T1-Tat2 complex binds to TAR1 approximately 10-fold less well than to TAR2. Significantly, the differences seen in Tat-CycT1 RNA binding activity correlate exactly with the ability of these two distinct Tat proteins to activate gene expression, thus explaining their differential abilities to activate heterologous LTR promoters (Bieniasz *et al.*, 1999b, Bieniasz *et al.*, 1999c). Furthermore, Tat2 is able to bind to cyclin T2A and cyclin T2B, alternative cyclin partners of CDK9, a property not possessed by Tat1, although these interactions are functionally abortive in that neither form can be recruited to TAR.

1.11.1.2 Tat recruits histone Acetyltransferase (HAT) proteins to the integrated HIV LTR promoter to mediate chromatin remodelling

Finally, more recent analyses has demonstrated that the optimal activity of Tat is further dictated by its association with a second class of cellular proteins, termed Tat-associated histone acetyltransferases (TAHs) (Benkirane *et al.*, 1998, Hottiger & Nabel, 1998, Marzio *et al.*, 1998, Kiernan *et al.*, 1999, Deng *et al.*, 2001, Hsia & Shi, 2002, Kino *et al.*, 2002). As previously described in section 1.10.1 the process

of transcription from an integrated promoter requires the rearrangement of the repressive chromatin structure in which it is bound (Kadonaga, 1998). Analyses of the chromatin structure of several stably integrated HIV-1 templates, using nuclease hypersensitivity and restriction endonuclease accessibility, indicate that the LTR is incorporated into two distinct nucleosomal regions, separated by a nuclease-hypersensitive region containing the TATA box and SP1/NF-κB binding sites (Verdin *et al.*, 1993, El Kharroubi *et al.*, 1998). When packaged into this repressive structure the HIV LTR is rendered transcriptionally silent in the absence of any stimuli (Verdin, 1991, Verdin *et al.*, 1993, Van Lint *et al.*, 1996, Widlak *et al.*, 1997, Benkirane *et al.*, 1998, Marzio *et al.*, 1998, Deng *et al.*, 2001, He *et al.*, 2002). The HIV Tat protein has been found to form a ternary complex with two highly homologous histone acetyltransferases, CBP/p300 and P/CAF, both *in vivo* and *in vitro* and in doing so targets them to the viral promoter (Hottiger *et al.*, 1998, Kiernan *et al.*, 1999). Following their recruitment CBP/p300 and P/CAF function to acetylate the histone proteins around which the promoter DNA is wound. This process results in the rearrangement of the chromatin structure encompassing the HIV LTR promoter such that it can be accessed by transcription factors, allowing the process of transcription to begin (Hottiger *et al.*, 1998, Kiernan *et al.*, 1999). In effect, the Tat protein acts as the trigger that de-represses the chromatinised LTR promoter and allows it to function.

As the number of cellular cofactors that are found to associate with HIV Tat *in vivo* grows (Brady *et al.*, 1999, Deng *et al.*, 2001, Ping *et al.*, 2001, Bharucha *et al.*, 2002, Kino *et al.*, 2002), the mechanisms by which Tat exerts its roles within HIV replication are being revealed and appear increasingly complex (Marcello *et al.*, 2001).

1.11.2. Effects of Tat on expression of heterologous viral and cellular genes

In addition to the role that the Tat protein plays in elongation enhancement, transcriptional synergy, and chromatin rearrangement, several investigators have shown that Tat functions to affect expression of selected viral and cellular genes in both infected and uninfected cells, stimulates cell growth, and induces apoptosis and T cell anergy. Most notably it has been demonstrated that Tat can activate the expression of several cytokines, including transforming growth factor-β, TNF, IL-2, and IL-6 (Chang *et al.*, 1995). In contrast, Tat has been implicated in the downregulation of the MHC class 1 promoter (Howcroft *et al.*, 1993), however this effect was found to be transient and limited (Verhoef *et al.*, 1998). In addition, Tat decreases the production of C-C chemokines that inhibit HIV-1 replication (Zagury *et al.*, 1998), and increases the expression of chemokine receptors that mediate infection by HIV-1 (Huang *et al.*, 1998a).

Mammalian cells have been shown to produce exogenous Tat, which is subsequently endocytosed by neighbouring cells and transported to the nucleus. In this way, Tat is able to exert its pleiotrophic effects upon cells that may not even be infected with HIV. HIV-1 Tat has been shown to stimulate growth of Kaposi's sarcoma (KS) cells and promotes adhesion of KS and normal vascular cells (Barillari *et al.*, 1993, Ensoli *et al.*, 1993, Albini *et al.*, 1995). In contrast, neither HIV-2 nor SIV Tat proteins have been associated with KS lesions.

Finally, a number of reports have indicated that Tat may play a role in dysregulation of the immune system. A study by Viscidi *et al* demonstrated that antigen-induced lymphocyte proliferation can be inhibited by the Tat protein (Viscidi *et al.*, 1989). In addition, Zagury and co-workers have recently reported that Tat may induce the generation of suppressor cells in populations of uninfected T cells, that may downregulate the immune response (Zagury *et al.*, 1998). Furthermore high anti-Tat antibody titres have been correlated with long term survival of HIV-1 infection, while Tat-specific CTL frequencies inversely correlate with progression to AIDS (Reiss *et al.*, 1991, van Baalen *et al.*, 1997). Taken together these findings support the notion that the Tat protein itself may play a role in HIV pathogenesis.

1.12 Biological relevance of the HIV-2 Long Terminal Repeat

The pathogenesis of HIV is closely related to plasma viral loads which are in turn influenced by the replicative capacity of the infecting virus population; differences in replicative capacities and plasma viral loads have been clearly associated with differences in rates of disease progression in HIV infection (Connor & Ho, 1994, Ho *et al.,* 1995, Wei *et al.,* 1995, Mellors *et al.,* 1996, O'Brien *et al.,* 1996, de Wolf *et al.,* 1997, Mellors *et al.,* 1997, Stein *et al.,* 1997, Berry *et al.,* 1998, Andersson *et al.,* 2000, Ariyoshi *et al.,* 2000, Popper *et al.,* 2000, Berry *et al.,* 2002, Alabi *et al.,* 2003). It has been suggested that the prolonged asymptomatic phase and freedom from clinical illness experienced by the majority of HIV-2 infected individuals results from a lower level of virus production during infection than that observed in either HIV-1 infection or HIV-2 infected individuals displaying relatively rapid progression to disease (Berry *et al.,* 1998, Andersson *et al.,* 2000, Ariyoshi *et al.,* 2000, Popper *et al.,* 2000, Berry *et al.,* 2002). It has been suggested that the level of virus production maybe lower as a result of more effective immune control of infection or results from a viral population of lower replicative capacity (Albert *et al.,* 1990, Tong-Starksen *et al.,* 1990, Hannibal *et al.,* 1993, Akimoto *et al.,* 1998, Grassly *et al.,* 1998, Sekigawa *et al.,* 1998, Reeves *et al.,* 1999, Kokkotou *et al.,* 2000).

The rate of virus production within an infected individual is determined in part by the potential of the viral LTR to support the transcription of the virally encoded genes utilising the *cis*-acting sequences in the promoter that interact with the cellular and viral transcription factors. Therefore, the sequences of the HIV-2 LTR have the potential to influence the replicative capacity of the viral population and thus the pathogenesis of HIV-2 infection.

In support of a role for the LTR in the control of replicative capacity, analyses of cultured HIV-1 isolates have revealed that increased replicative capacity in permissive cell lines can be attributed to variant LTR sequences. For example, Golub *et al* showed that a 24bp insertion 5' to the NF-κB sites in the HIV-1 LTR was responsible for a 3-fold increase in the replicative capacity of an HIV-1 isolate (Golub *et al.,* 1990). Furthermore, Zhang and co-workers have described HIV-1 variants with deletions to the NFAT sites of the LTR, these variants were found to replicate faster and were more cytopathic *in vitro* than parental strains (Zhang *et al.,* 1997a). More recently, duplication of the SP1 region through prolonged culturing of an attenuated HIV-1 subtype B virus was reported, this LTR variant was a stronger promoter with six SP1 sites and yielded a fitter virus (Berkout *et al.,* 1999).

In addition, several groups have provided evidence of enhanced replicative capacity associated with genetic variation within the promoter regions of naturally occurring HIV-1 isolates. Koken *et al* screened for 5' HIV-1 LTR variants from 17 patients and identified an LTR variant with a fourth SP1 site that could outgrow an isogenic construct with three SP1 sites, as well as length variants with a 5'-CTG-3' motif insertion 5' of NF-κB sites that could confer a marginal decrease in transcription and replicative capacity (Koken *et al.*, 1992).

In evaluating group M HIV-1 LTR sequences from viral subtypes A through G Jeeninga *et al* have recently reported notable differences in the replication of the subtypes which may be due to genetic variation in the LTR promoter (Jeeninga *et al.*, 2000). The LTR taken from a subtype E isolate displayed three-fold higher basal activity than a subtype B LTR; when introduced into the subtype B isolate the LTR core promoter elements of subtype E conferred increased replication rates.

Most interestingly, Fang and co-workers have identified specific sequence changes within the LTR of virus taken from an HIV-1-infected individual that correlate with the transition from a long-term non-progressive infection to a progressive infection (Fang *et al.*, 2001). This study suggests that a large deletion in one of the SP1 binding sites of the HIV-1 LTR played a role in decreasing virus replication within this patient, resulting in low viral loads and long-term nonprogressive infection. Extensive functional and sequence analyses of the SP1 promoter region during the transition from nonprogressive to progressive infection indicated that the transition to high viral loads and rapidly progressive infection within this patient were related to the proliferation of a viral population that had intact SP1 and promoter sequences.

Despite the evidence supporting a determining role for the LTR in HIV replicative capacity and pathogenesis, several investigations comparing the sequence and function of the LTRs from a number of different LTNP and RP HIV-1 infected cohorts have failed to identify any temporal enrichment of particular LTR sequences during infection (Delassus *et al.*, 1991, Delassus *et al.*, 1992), or a general correlation between promoter activity and disease state (Estable *et al.*, 1996, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Quiñones-Mateu *et al.*, 1998, Visco-Comandini *et al.*, 1999, Gómez-Román *et al.*, 2000). Nevertheless, comparative functional and sequence analyses have yet to be performed upon cohorts of HIV-2 infected patients exhibiting different rates of disease progression.

1.13 Project rationale and objectives

Since the majority of research into the pathogenic significance of the HIV LTR has centred on samples taken from HIV-1 infected individuals, characterisation of the genotypic and phenotypic properties of uncultured HIV-2 LTR variants is required. Molecular and biochemical analyses of the region upstream of the HIV-2 LTR core promoter have identified several modulatory *cis*-acting elements that are conserved amongst HIV-2/SIV isolates, but distinct in sequence and function from those elements located in the corresponding region of the HIV-1 LTR (Markovitz *et al.*, 1990, Markovitz *et al.*, 1992, Clark *et al.*, 1995). These elements have been shown to regulate differentially expression from the HIV-2 LTR in comparison to the HIV-1 LTR (Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994). It has been suggested that the difference in regulation of these two promoters may, by virtue of the relationship between promoter activity, replicative capacity, and disease prognosis, provide a mechanism by which the difference in pathogenicity of the two viruses is determined (Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Faulkner *et al.*, 2001). Similarly, it is possible that differences or changes in the function of the HIV-2 LTR and/or its regulation could provide a mechanism by which the pathogenicity of HIV-2 the subsequent rate of disease progression is virally determined. It is possible that the lower level of virus production and extended asymptomatic phase observed within the majority of HIV-2 infections results from an infecting population that has attenuated promoter function in comparison to the populations found within HIV-2 infected individuals exhibiting rapid progression to disease or HIV-1-infected individuals. An investigation into the function of naturally occurring and isolated HIV-2 LTR genotypes derived from individuals displaying a variety of disease progression status would enable an assessment of the role, if any, that the HIV-2 LTR plays in the determination of disease progression rate and HIV-2 pathogenicity.

While the majority of countries in the Western world can treat and control HIV infection using highly active antiretroviral drug therapy (HAART), the continents most affected by HIV, Africa and Asia, are the least able to afford, or administer, HAART effectively. At present in many of these countries, funding is only available to treat STDs in an effort to reduce the spread of HIV infection. Therefore, development of a vaccine offers the most likely and affordable means to prevent and control the spread of HIV in these areas. While the provision of a preventative vaccine is required, a vaccine based prophylactic treatment for those already infected is also a major goal. In this vein characterisation of the determinants of HIV pathogenesis may help guide the development of treatments or vaccines to delay the onset of AIDS in both HIV-2 and HIV-1 infections. In addition, details of the mechanisms that determine the prolonged

asymptomatic period observed in HIV-2 infection prior to the onset of AIDS are extremely important to the development of strategies aimed at extending the asymptomatic phase of HIV-1 infection. In summary, the HIV-2 infection provides a most valuable *in vivo* and *ex vivo* comparator with HIV-1 in terms of pathogenesis, control of viral replication and viral phylogeny.

The aims of this project are two-fold. Firstly, to develop and optimise efficient PCR amplification, cloning, transfection, and luciferase reporter assay procedures for the rapid and reproducible recovery and measurement of HIV-2 LTR activity from high and very low copy number uncultured PBMC-derived DNA samples alike. Using a dual luciferase reporter assay we aim to accurately and reproducibly study the function of isolate and naturally occurring HIV-2 LTR genotypes derived from clinically characterised individuals exhibiting different disease progression statuses, at both basal and activated levels, within two biologically relevant cell lines. Secondly, amplified HIV-2 LTRs that have been functionally characterised will be sequenced in order to assess genetic heterogeneity. This will provide an assessment of the biological significance of functional and genetic diversity in the HIV-2 LTR with respect to viral pathogenicity and clinical disease progression.

Chapter 2.
Materials and Methods.

2.1 Materials

2.1.1 HIV-2 Infected Material/Sources

2.1.1.1 HIV-2 Isolates

CBL Series - Chester Beatty Laboratory (CBL) viral isolates (MRC ADP) 20-24. Isolated from Gambian patients exhibiting a differing range of *in vitro* phenotypes that correlate with the patients' clinical status. The five Gambian HIV-2 strains (CBL 20-24) were isolated at the Institute for Cancer Research, Chester Beatty Laboratories, London (Schultz *et al.*, 1990) and range from a highly cytopathic and fast replicating virus from a patient who died of AIDS (CBL 20), to noncytopathic slow replicating virus from an asymptomatic patient (CBL 24) (Schulz *et al.*, 1990). Isolates were cultured in H9 or C8166 cells at 37°C in 5% CO_2 in category III facilities. DNA was extracted from each isolate using the method described in section 2.2.2.1a, all procedures were carried out in class I cabinets using gamma irradiated sterile plasticware (Falcon) by colleagues at the Chester Beatty Laboratories.

2.1.1.2 HIV-2 Clinical Material - The Gambian Cohort-

Several PBMC samples from HIV-2 seropositive patients attending the STD clinic at the MRC hospital in Fajara, The Gambia, were obtained for the purpose of this study. The samples obtained were selected from a clinically defined longitudinal cohort consisting of patients who had entered the clinic and been characterised by clinical and laboratory examination (performed by colleagues at the MRC laboratories, Fajara, The Gambia). Patients recruited into this study exhibited either rapid progression to AIDS in comparison to normal clinical cases of HIV-2, or long term non progression where the patient was asymptomatic and has to date not developed any symptoms associated with progression to AIDS. Patients were sampled at 6 monthly intervals by colleagues at the MRC laboratories, Fajara, The Gambia, and disease progression was defined on the basis of CD4% decline over the period of clinical follow-up. Non-progressors were classified as having stable CD4% for more than 56 months, while rapid progressor patients were those with CD4% declining from greater than 20% to less than 13% within 20 months. In addition to CD4 status, patients were characterised according to plasma viral and proviral loads (determined by Dr. N. Berry, NIBSC, see section 2.2.1). Proviral DNA was extracted using the method described in section 2.2.2.1b, and stored at -20°C.

2.1.2 Bacterial Strains

Bacterial strains were grown at 37°C aerobically in 2xTY media (Appendix I). Bacterial manipulations were carried out using aseptic techniques with sterile glass and plasticware.

Strain	Use
SCS-1	Plasmid purification, calcium chloride transformation and ampicillin resistance screening, *recA*$^+$.
XL-1 Blue	Plasmid purification, calcium chloride transformation and ampicillin resistance screening.
Epicurian coli® SCS-1	Supercompetent *recA* deficient cells harbouring an uncharacteristic mutation that improves transformation efficiency. Reduced homologous recombination of propagated DNA. (Stratagene)

Table 2.1 Bacterial strains

2.1.3 Mammalian Cell Lines

All mammalian cell lines were grown at 37°C in 5% CO_2 humid atmosphere using sterile tissue grade reagents and gamma irradiated plasticware (Falcon). Tissue culture procedures were carried out in a class II laminar flow cabinet.

Cell Line	Source	Growth Medium	Description
293	Dr. J. McKeating	DMEM + 10% FCS	Human embryonic kidney cell, transformed with sheared adenovirus type 5 DNA
Jurkat Clone E6-1	MRC ARP	RPMI 1640 + 10% FCS	Human T-cell leukaemia cell line, clone of Jurkat-FHCRC.
THP-1	MRC ARP	RPMI 1640 + 10% FCS	Human acute monocytic leukaemia cell line derived from macrophage lineage (CD14+; CD15+).

Table 2.2 Mammalian cell-lines

2.1.4 Plasmid Vectors

Plasmid stocks were stored resuspended in sterile dH_2O at -20°C, as *E.Coli* glycerol stocks also stored at -20°C and agar plate grids at 4°C.

Name	Source	Description
PGL3-Enhancer (pGL3E)	Promega Ltd	Mammalian reporter vector containing a firefly luciferase reporter gene downstream of a polylinker, permitting insertion of upstream promoter elements. pGL3E contains an SV40 enhancer located downstream of the luciferase gene and the poly(A) signal resulting in transcription of *luc+* at higher levels.
pGL3-Control (pGL3C)	Promega Ltd	Mammalian reporter vector containing a firefly luciferase reporter gene downstream of an SV40 promoter. pGL3C also contains an SV40 enhancer located downstream of the luciferase gene and a poly(A) signal.
pRL-TK	Promega Ltd	*Renilla* luciferase internal control vector containing a cDNA (R*luc*) encoding *Renilla* luciferase downstream of the herpes simplex virus thymidine kinase (HSV-TK) promoter. Provides low to moderate levels of *Renilla* luciferase expression in cotransfected mammalian cells.
pRL-CMV	Promega Ltd	*Renilla* luciferase internal control vector containing a cDNA (R*luc*) encoding *Renilla* luciferase downstream of the cytomegalovirus (CMV) enhancer and early promoter elements. Provides high-level *Renilla* luciferase expression in cotransfected mammalian cells.
pRL-SV40	Promega Ltd	*Renilla* luciferase control vector containing a cDNA (R*luc*) encoding *Renilla* luciferase downstream of the SV40 enhancer and early promoter elements. Provides high-level *Renilla* luciferase expression in cotransfected mammalian cells.
RSVtat2	Dr. D. Markovitz	Mammalian expression vector containing an HIV-2 *tat* gene downstream of the Rous Sarcoma Virus promoter.
$pROD_{10}$	MRC ARP	Infectious molecular clone of HIV-2, based on $HIV\text{-}2_{ROD}$

Table 2.3 Plasmid vectors

2.1.5 Enzymes

Enzymes were used according to manufactures instructions with regard to optimal working temperatures and buffer requirements. Digestions requiring more than one restriction endonuclease were carried out in buffers that allowed > 70% activity for each enzyme, according to the manufacturers supplied data.

Enzyme	**Manufacturer**
RNAse	Sigma-Aldrich Company Ltd, Dorset, UK.
Pfu Turbo™ DNA Polymerase	Stratagene Europe Amsterdam, The Netherlands.
Restriction Endonucleases	Boehringer-Mannheim, Lewes, UK. Promega Ltd, Southampton, UK. New England Biolabs (UK) Ltd, Hertfordshire, UK.
Shrimp Alkaline Phosphatase	United States Biochemicals
T4 DNA Ligase	Promega Ltd. Southampton, UK
Thermo Sequenase DNA Polymerase	Amersham pharmacia biotech, Buckinghamshire, UK.

Table 2.4 Enzymes and manufacturers used

2.1.6 Oligonucleotides

The design of oligonucleotides is described for each polymerase chain reaction individually. All primers were obtained from Oswel DNA Service, Southampton, UK.

2.2 Methods

2.2.1 Patient Characterisation.

2.2.1.1 $CD4^+$ cell Status

An estimation of the percentage of CD4 positive cells from HIV-2 positive whole heparinised blood was obtained using fluorescent activated cell sorting (Becton-Dickinson) after staining with an anti-CD4 monoclonal antibody (performed by colleagues at the MRC laboratories Fajara, The Gambia). Light microscopy was used to estimate the total white cell count and lymphocyte percentage. CD4 percentages (CD4%) were established as being a more robust measurement than absolute CD4 counts and were therefore used in studies in The Gambia.

2.2.1.2 DNA viral load

HIV-2 proviral DNA load was assessed by Dr N Berry, NIBSC. Peripheral blood mononuclear cells were prepared from heparinised whole blood using Lymphoprep™ (Micromed Diagnostics). A PCR compatible extraction buffer and proteinase K was used to extract cellular DNA. The concentration of cellular DNA was measured by fluorometry using a TK100 fluorormeter (Hoefer Scientific Instruments) and fluorescent staining with Hoescht 33258 dye, DNA concentration was adjusted to 60μg/ml. HIV-2 proviral load was quantified by nested PCR amplification of a highly conserved region within the long terminal repeat, followed by a solid-phase radiometric incorporation assay (Berry *et al.*, 1994).

2.2.1.3 RNA viral load

RNA viral load was determined by Dr N Berry, NIBSC. RNA was prepared from 200μl patient plasma using a silica-based guanidinium isothiocyanate extraction method. RNA samples were subjected to cDNA reaction using Superscript II reverse transcriptase, (Gibco BRL) cDNA copy number was determined by limit dilution analysis and quantitative RNA estimations by chemiluminescence. The lowest detection range for this assay system was calculated to be 200 copies of HIV-2 RNA per ml of plasma (Berry *et al.*, 1998).

2.2.2 DNA analysis and manipulation

2.2.2.1 Extraction and quantification of cellular DNA.

2.2.2.1a HIV-2 Isolates.

DNA was extracted from a subset of five virus isolates from the series of isolates collectively known as the Chester Beatty Laboratory (CBL) series. The five Gambian HIV-2 strains (CBL 20-24) isolated at the Institute for Cancer Research, Chester Beatty Laboratories, London (Schultz *et al.*, 1990), were established in continuous cell lines. Each isolate was cultured at 37°C in 5% CO_2, all procedures were carried out in class I cabinets under category III facilities using gamma irradiated sterile plasticware (Falcon). Infected cells were thawed rapidly at 37°C, pelleted by centrifugation, and washed in PBS (Appendix I). Cell pellets were resuspended in a lysis buffer containing 0.1% SDS and 10μg/ml Proteinase K. Virus inactivation and cell lysis was achieved by incubation at 65°C for 2 hours. DNA was extracted from lysed cells using a phenol-chloroform, ethanol precipitation method. 0.5mls of phenol:chloroform:isoamyl alcohol (25:24:1) was mixed with lysed cells in a polypropylene tube, until an emulsion formed. Following centrifugation the aqueous phase was removed, a further chloroform extraction was performed to remove trace amounts of phenol. DNA was precipitated by the addition of 2 volumes of ice cold absolute alcohol and 1/10th 3M sodium acetate, pH4, followed by incubation at -20°C overnight. Precipitated DNA was pelleted in a bench microfuge at 13000rpm for 30 minutes, washed with 70% ethanol, resuspended in sterile TE buffer (Appendix I) and stored at -20°C. All procedures were performed by colleagues at the Chester Beatty Laboratories, London.

2.2.2.1b Patient Peripheral Blood Mononuclear Cells

Cellular DNA was prepared from aliquots of peripheral blood mononuclear cells using a QIAamp Midi Blood Kit (Qiagen) according to the manufacturers instructions. PBMCs had been isolated using Lymphoprep (Mycomed Ltd) by colleagues in The Gambia, transported to London on dry ice, and stored at -70°C in freezing medium (Appendix I). Each aliquot of frozen cells was thawed rapidly in a 37°C water bath, equilibrated to room temperature, and treated according to the kit instructions (Qiagen). DNA was eluted in sterile distilled water and stored at -20°C.

2.2.2.2 Measurement of DNA concentration by fluorometry

The concentrations of DNA prepared from virus isolate, PBMC, or bacterial plasmid transformation was measured by fluorometry. DNA concentrations were measured by a 96 well plate format fluorescent dye-

binding assay using a Fluoroskan Ascent FL (Fluorometer/Luminometer, Life Sciences International (UK) Ltd). The assay accurately measured a DNA concentration range from 5 to 1600μg/ml (equivalent to the DNA from 10^6 to $3x10^8$ cells in 1 ml). The fluorometer was standardised using seven calf thymus DNA standards (32ng/ml to 32000ng/ml), and measured excitation and emission wavelengths of 355nm and 460nm respectively. An aliquot of 150μl of standard or diluent (Appendix I) was dispensed into black flat-bottom microtitre wells (Nunc) as shown in Figure 2.1, pg 108, followed by 3μl of sample DNA to the test wells. An aliquot of Hoecht 33258 dye was diluted 1:200 with assay diluent (Appendix I) such that each well containing standard or test DNA would receive 50μl of the diluted dye solution. The diluted Hoecht was dispensed into the wells and the plate was placed in a rotating plate shaker for approximately ten seconds and placed in the dark for five minutes before reading. The plate was then loaded into the fluorometer and DNA concentrations were measured and calculated using Labsystems Fluoroscan Ascent software.

Blank	Std 4	T1	T5	T9	T13	T17	T21	T25	T29	T33	T37
Blank	Std 4	T1	T5	T9	T13	T17	T21	T25	T29	T33	T37
Std 1	Std 5	T2	T6	T10	T14	T18	T22	T26	T30	T34	T38
Std 1	Std 5	T2	T6	T10	T14	T18	T22	T26	T30	T34	T38
Std 2	Std 6	T3	T7	T11	T15	T19	T23	T27	T31	T35	T39
Std 2	Std 6	T3	T7	T11	T15	T19	T23	T27	T31	T35	T39
Std 3	Std 7	T4	T8	T12	T16	T20	T24	T28	T32	T36	T40
Std 3	Std 7	T4	T8	T12	T16	T20	T24	T28	T32	T36	T40

Blank = diluent only, T= Test, Std = Standard.

Figure 2.1 Diagram of the sample positions within the black flat-bottom microtitre plate (Nunc).

2.2.2.3 PCR Amplification of HIV-2 Long Terminal Repeat.

A nested PCR method was developed for the amplification of the 3' LTR region by targeting HIV-2 proviral DNA from clinical material, *in vitro* infected-cell provirus, and HIV-2 molecular clones. The PCR method described below was adapted from the original PCR methods for *in vitro* amplification of specific DNA sequences (Mullis & Faloona, 1987). The initial nested PCR protocol described was further optimised for HIV-2 LTR amplification from patient material exhibiting low proviral copy

number (<200 copies/ml). The resulting PCR was limiting dilution sensitive and capable therefore of amplifying from a single molecule of HIV-2.

2.2.2.3a Oligonucleotide design and synthesis

Oligonucleotide primers were designed to target the U3, R, and part of the U5 region of the HIV-2 LTR. Primer sequences were determined following reference to published HIV-2 sequences recorded in the Los Alamos Database. Sequences were analysed using Microgenie software (Beckman Instruments Ltd). All primers were designed to be between 18 and 25 nucleotides in length and have as near possible a 50% G/C content. Oligonucleotides were synthesised by Oswel DNA Service (University of Southampton). The theoretical melting temperature (Tm) for each primer was calculated according to:

$$Tm = 4\,(G+C) + (A+T) \text{ minus } 5^{\circ}C$$

Second round primers were designed at positions +8934 and +180 according to the numbering of HIV-2 ROD, GenBank accession number M15390. LTR regions outside of these sites had been shown to be superfluous to LTR function (Guyader *et al.*, 1987, Arya & Gallo, 1988, Markovitz *et al.*, 1990, Tong-Starksen *et al.*, 1990, Arya, 1991, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995). The second round cloning primers were designed to contain 5' restriction endonuclease sites, which matched sites, found within the polylinker region of firefly luciferase reporter vector pGL3E (Promega Ltd). The primers used in this study are listed in Table 2.5, pg 110.

Primers for nested amplification of HIV-2 Long Terminal Repeat (LTR).

Palindromic restriction endonuclease sequences are shown underlined, ACGCGT = *Mlu I* and AGATCT = *Bgl II*

Primer Name	**Orientation**	**Location (nt)[a]**	**Sequence**
Z4544	Sense Outer	8557-8582	5' ATGGGTGCGAGTGGATCCAAGAAGC 3'
Y6461	Antisense Outer	287-313	5' TTCGGGCGCCAACCTGCTAGGGATTTT 3'
T5169	Sense Inner	8934-8949	5' GCCCGTACGCGTGGGGGGACTGGAAGGG 3'
T5170	Antisense Outer	180-197	5' GGCCAGATCTAGAGATGGGARCACACAC 3'

Table 2.5 Oligonucleotide primer sequences for nested PCR amplification of HIV-2 Long Terminal Repeat. [a] Sequences numbered according to HIV-2 ROD, GenBank accession number M15390.

2.2.2.3b Nested PCR amplification of HIV-2 LTR from virus isolate, molecular clone and high copy number clinical material

The HIV-2 LTR was amplified from patient provirus, isolate-derived proviral DNA, and molecular clone plasmid DNA, by a two round nested PCR using specific sense and antisense oligonucleotide primers (see Table 2.5, pg110). First round reaction conditions were 100mM KCl, 100mM $(NH_4)_2SO_4$, 200mM Tris-HCl (pH 8.8), 20mM $MgSO_4$, 1% Triton X-100 and 1mg/ml BSA), (10x cloned *Pyrococcus furiosus* (*Pfu*) Turbo buffer (Stratagene)), 200µM of each dNTP (A, C, G, and T), 5 units cloned *Pfu* Turbo DNA polymerase (Stratagene) and 0.1µM of each outer primer, in a reaction volume of 50µl. The reaction mixture was overlaid with 30µl of light white mineral oil. First round cycling conditions were preceded by a pre-melting step of 94°C for 4 mins for 1 cycle, followed by 94°C 1 min, 50°C 1 min, 72°C 1.5mins for 35 cycles and one cycle at 72°C for 7 mins then cooled and maintained at 1°C.

For second round PCR amplification 2µl of first round PCR product was transferred into 50µl of the following second round reaction mixture, 100mM KCl, 100mM $(NH_4)_2SO_4$, 200mM Tris-HCl (pH 8.8), 20mM $MgSO_4$, 1% Triton X-100 and 1000µg/ml BSA), (10x cloned *Pyrococcus furiosus* (*Pfu*) Turbo buffer (Stratagene)), 200µM of each dNTP (A, C, G, and T), 5 units cloned *Pfu* Turbo DNA polymerase (Stratagene). 0.1µM of each second round cloning primer instead of the outer primers was added in a reaction volume of 50µl, the reaction mixture was then overlaid with 30µl of light white mineral oil. Second round cycling conditions: 94°C for 4 mins for 1 cycle, 94°C for 45 secs, 54°C for 45 secs, 72°C for 1 min for 30 cycles, followed by 72°C for 7 mins then cooled and maintained at 1°C. Amplification reactions were performed on a DNA thermal Robocyler (Stratagene).

Amplified PCR products were visualised by loading 10µl of second round reaction mixture onto a 1% agarose gel containing 0.5mg/ml ethidium bromide (Appendix I). The gel was electrophoresed at 100v for 30 mins in 1xTAE buffer (Appendix I), and photographed under ultra violet illumination.

2.2.2.4 PCR optimisation to single copy sensitivity

2.2.2.4a Robocycler™ determination of optimum PCR annealing temperature

Initial estimations of the optimal primer theoretical melting temperature (Tm) were made using the equation in section 2.2.2.3a. The experimental temperature optimum for the second round primer pair was determined using the Robocycler 480™ (Stratagene). In contrast to other thermocycling machines,

this machine contained four heating blocks independently thermally controlled, for each step of the PCR. A robotic arm moved the reaction tubes into each block (set at the required temperature) for the duration of each stage of the PCR cycle, as opposed to a single block determining the heating and cooling cycles. The block designated for the annealing reaction has the ability to heat either to a single temperature, or to a temperature gradient spanning the width of the block. The heating block can accommodate 40 reaction tubes in a 5 x 8 well format. This allows for the PCR annealing temperature optimisation of five samples over the range of eight different temperatures. The temperature gradient facility of the Robocycler 480 was utilised to determine the optimal annealing temperature of the primer pair used in the second round HIV-2 LTR PCR. Second round PCR conditions were designed with an annealing reaction temperature gradient that included the calculated theoretical melting temperature of each second round primer.

The first round PCR conditions matched those described in section 2.2.2.3b. The reaction conditions for the second round were 100mM KCl, 100mM $(NH_4)_2SO_4$, 200mM Tris-HCl (pH 8.8), 20mM $MgSO_4$, 1% Triton X-100 and 1mg/ml BSA), (10x cloned *Pyrococcus furiosus* (*Pfu*) Turbo buffer (Stratagene)), 200µM of each dNTP (A, C, G, and T), 5 units cloned *Pfu* Turbo DNA polymerase (Stratagene). 0.1µM of each second round cloning primer was added to a reaction volume of 50µl. Eight tubes were aliquoted with 50µl of second round reaction mixture. Two microlitre aliquots of first round PCR product were transferred to each of the eight tubes and each reaction mixture was overlaid with 30µl of light white mineral oil. The thermocycling programme for the second round was preceded by a pre-melting step of 94°C for 4 mins for 1 cycle, followed by 94°C for 45 seconds, 48°C-62°C in two degree increments (eight temperatures over eight wells of the block) for 45 seconds, 72°C for 1 min for 30 cycles, followed by 72°C for 7 mins then cooled and maintained at 1°C. PCR products were visualised as described in section 2.2.2.3b.

2.2.2.4b Determination of nested PCR sensitivity by limiting end-point dilution

The sensitivity of the second round primer pair was assessed by limiting end-point dilution. A first round reaction mixture was prepared as described in section 2.2.2.3b and distributed into ten sets of four tubes in the following manner, each tube ultimately containing 20µl. The first tube was aliquoted with 95µl of reaction mixture, to this tube, 5µl of sample DNA was added, and the mixture was vortexed for 5 seconds. The remaining three tubes of the first set of four were aliquoted with 20µl of reaction mix from the first tube, the final 20µl was then transferred into 80µl of reaction mixture in the first tube of the

second set of four. Twenty microlitre aliquots were transferred into the remaining tubes of the second set, this continued until 20µl was diluted in 60µl reaction mixture in the first tube of the final set. This method resulted in a dilution series of sample DNA of 1, 1/5, 1/25, 1/125, 1/625, and so on. Second round nested PCR reaction conditions matched those described in section 2.2.2.4a and were prepared using 2µl transfers from first round PCR reactions. Second round cycling conditions were as follows: 94°C for 4 mins for 1 cycle, 94°C for 45 secs, 50°C for 45 secs, 72°C for 1 min for 30 cycles, followed by 72°C for 7 mins then cooled and maintained at 1°C. Amplification reactions were performed on a DNA thermal Robocyler (Stratagene) and analysed by gel electrophoresis. Wells were scored PCR positive or negative and the frequency of zero (Fo) amplification products in each set of four tubes was scored. According to the zero term of the Poisson probability theory the average number of DNA molecules per tube (µ) can be assessed by inserting the number of non-amplifications per set into the Poisson equation:

$$\mu = -\ln Fo$$

The natural logarithm (ln) of the fraction of negative values (Fo) is proportional to the mean of the number of positive reactions (µ) for each group. Taking the dilution factor into account the number of DNA molecules per microlitre of test sample could be calculated providing a measurement of the copy number of HIV-2 molecules. The presence of a Poisson distribution end point indicates that the primer pairing in the nested PCR is capable of amplifying from a single molecule of DNA.

2.2.2.5 Amplification of HIV-2 LTR from low copy number patient material using a limiting dilution sensitive nested PCR.

The HIV-2 LTR was amplified from low copy number patient provirus by a two round nested PCR which had be shown to be limiting dilution sensitive. The specific sense and antisense oligonucleotide primers were the same pairs as those described in section 2.2.2.3 (see Table 2.5, pg 110). First round reaction conditions were 100mM KCl, 100mM $(NH_4)_2SO_4$, 200mM Tris-HCl (pH 8.8), 20mM $MgSO_4$, 1% Triton X-100 and 1mg/ml BSA), (10x cloned *Pyrococcus furiosus* (*Pfu*) Turbo buffer (Stratagene)), 200µM of each dNTP (A, C, G, and T), 5 units cloned *Pfu* Turbo DNA polymerase (Stratagene) and 0.1µM of each outer primer, in a reaction volume of 50µl. The reaction mixture was overlaid with 30µl of light white mineral oil. First round cycling conditions were preceded by a pre-melting step of 94°C for 4 mins for 1 cycle, followed by 94°C 1 min, 50°C 1 min, 72°C 1.5mins for 35 cycles and one cycle at 72°C for 7 mins then cooled and maintained at 1°C.

For second round PCR amplification 2µl of first round PCR product was transferred into 50µl of the following second round reaction mixture, 100mM KCl, 100mM $(NH_4)_2SO_4$, 200mM Tris-HCl (pH 8.8), 20mM $MgSO_4$, 1% Triton X-100 and 1mg/ml BSA), (10x cloned *Pyrococcus furiosus* (*Pfu*) Turbo buffer (Stratagene)), 200µM of each dNTP (A, C, G, and T), 5 units cloned *Pfu* Turbo DNA polymerase (Stratagene). 0.1µM of each second round cloning primer instead of the outer primers was added in a reaction volume of 50µl, the reaction mixture was then overlaid with 30µl of light white mineral oil. Second round cycling conditions: 94°C for 4 mins for 1 cycle, 94°C for 45 secs, 50°C for 45 secs, 72°C for 1 min for 30 cycles, followed by 72°C for 7 mins then cooled and maintained at 1°C. Amplification reactions were performed on a DNA thermal Robocyler (Stratagene).

Amplified PCR products were visualised as before, by loading 10µl of second round reaction mixture onto a 1% agarose gel containing 0.5mg/ml ethidium bromide (Appendix I). The gel was electrophoresed at 100v for 30 mins in 1xTAE (Appendix I) buffer, and photographed under ultra violet illumination.

2.2.2.6 Isolation and purification of PCR-amplified products

PCR products were isolated using agarose gel electrophoresis, this procedure removed residual enzyme and excess primers contained within the PCR reaction mixture. The 764bp LTR band was resolved by loading 40µl of a 50µl PCR onto a 1% agarose gel containing 0.5mg/ml ethidium bromide (Appendix I); the gel was electrophoresed at 100V for 1 hour. The LTR band was visualised using low level UV irradiation and excised from the gel using a sterile scalpel blade. DNA was extracted from the agarose slice using the GeneClean™ (Bio 101, La Jolla, CA) silica resin procedure or the QIAquick™ (Qiagen) column extraction method according to manufacturers instructions. Purified PCR products were resuspended in 20µl of sterile dH_2O and stored at -20°C.

2.2.3 Cloning of HIV-2 LTR PCR products

2.2.3.1 Restriction endonuclease digestion of amplified LTR PCR products

To enable efficient cloning of PCR amplified LTRs into the firefly luciferase reporter vector pGL3E (Promega Ltd), PCR product were digested with restriction endonucleases. The flanking restriction sites

incorporated into the PCR products by the inner cloning primers matched two sites contained within the polylinker region of pGL3E. Once digested the PCR product contained complementary ends to those of the digested linearised reporter vector, enabling the formation of a complete double-stranded circularised plasmid containing an LTR insert. Digestion reactions were carried out at 37°C for 2 hours in 20µl volumes. 10µl of purified PCR product: 0.5 Units *Mlu I*, 0.5 Units *Bgl II*, 2µl 10x Buffer H (Appendix I)(Boehringer Manheim), was made to the total reaction volume with dH_2O. Following incubation, restriction endonucleases were removed from this mixture using Strataclean™ Resin according to the manufacturers instructions (Stratagene) and stored at -20°C.

2.2.3.2 Linearisation of Firefly Luciferase Reporter Plasmid pGL3E (Promega Ltd)

To enable efficient cloning of PCR amplified LTRs into the firefly luciferase containing reporter plasmid it was necessary to linearise the vector using *MluI* and *BglII* sites found within the polylinker region, matching those designed into the second round primers. 1µg of pGL3E was added to 20 units *MluI* (Boehringer Mannhiem) in 1x Buffer H (Appendix I) (Boehringer Mannhiem) and incubated for 1 hour at 37°C; residual enzyme was removed with 5µl Strataclean™ resin (Stratagene). This was followed by addition of 20 units *BglII* (Boehringer Mannhiem) and incubation at 37°C for a further 1 hour.

2.2.3.3 Dephosphorylation of Linearised Firefly Luciferase Plasmid

Following restriction endonuclease digestion linearised pGL3E was dephosphorylated to prevent self-ligation. 1 Unit of Shrimp Alkaline Phosphatase (United States Biochemicals) was added to the digestion reaction mixture and incubated at 37°C for 15 mins. Inactivation of Phosphatase was achieved by a further incubation for 15 mins at 65°C.

2.2.3.4 Cloning of PCR amplified LTR into linearised pGL3E

Restriction endonuclease digested PCR products were ligated into linearised pGL3E using 4 Units of T4 DNA ligase (Promega Ltd) in a total ligation mixture volume of 20µl. The mixture contained the appropriate dilution of manufacturer-supplied buffer (Appendix I) and a vector to insert ratio of 1:3, 1:5, or 1:10. Ligation reactions were carried out at 12°C overnight. Protocols using the vector to insert ratio of 1:10 routinely gave optimum ligation efficiencies of 70-80%.

2.2.3.5 Preparation and transformation of competent E.coli

A vented conical flask containing 50mls of 2xTY broth (Appendix I) was inoculated with 500µl of an overnight culture of SCS-1, or XL-1Blue *E.coli*. Cultures were grown for 3 hours at 37°C in an orbital shaker to ensure complete aeration. Cultures were chilled on ice for 10mins and cells were pelleted in a prechilled centrifuge at 4000 rpm at 4°C for 10 mins. Cells were resuspended in 25ml ice-cold 0.1M Calcium Chloride (Appendix 1) and incubated on ice for 1 hour. Following incubation, cells were repelleted as before and gently resuspended with ice-cold 0.1M Calcium Chloride to a final volume of 2.5ml. Competent cells were either transformed immediately or stored at 4°C packed in ice for up to 5 days.

Transformations were carried out using 100µl of competent cells and up to 1µg of DNA. DNA was added to round-bottomed transformation tubes (Falcon) and overlaid with competent cells using a cut 200µl pipette tip; mixtures were swirled gently and incubated on ice for 1 hour. Cells were heat-shocked at 42°C in a water bath for 90 seconds and put back on ice for 10 mins prior to plating out on selection media containing the appropriate antibiotic (Appendix I). Competency of cells was tested by transforming 1ng of a commercially prepared antibiotic resistance plasmid such as pGL3E (Promega Ltd).

2.2.3.6 Transformation of commercially available competent cells

Epicurian coli® SCS-1 supercompetent cells (Stratagene) were transformed according to manufacturers instructions, and plated out onto selection media containing the appropriate antibiotic.

2.2.3.7 PCR Screening of bacterial transformants

Ampicillin resistant colonies were screened for the presence of the LTR sequence by picking colonies, using a flamed loop, directly into an eppendorf containing 20µl of PCR mix used in second round PCR amplification reactions (section 2.2.2.5). Cycling conditions were the same as those used in the second round PCR reaction although only 25 cycles of the reaction was necessary. Primers were able to reach their target as the first cycle of heating lysed the bacterial cells. Colonies, which contained a plasmid with an LTR insert, were identified using 1% agarose gel electrophoresis. Cloning efficiency was calculated by dividing the number of positive clones by the total number of colonies tested. Positive colonies were grown up as overnight cultures in 3mls 2x TY broth (Appendix I) supplemented with ampicillin at 100mg/ml and were placed in a rotary shaker at 37°C (section 2.2.3.8).

2.2.3.8 Growth of Bacterial Strains

For plasmid DNA purification bacterial strains were grown aerobically in 2xTY broth (Appendix 1) at 37°C in an orbital shaker for 16 hours. Bacterial screening and colony isolation was carried out by growing cultures on 2xTY 1.5% agarose plates (Appendix I); bacterial transformants were selected by the addition of ampicillin at 100mg/ml. Bacteria were stored as glycerol stocks (500µl overnight culture plus 500µl autoclaved glycerol) at -20°C, or sealed agar plate grids at 4°C.

2.2.3.9 Isolation of plasmid DNA

Plasmid DNA was extracted from either 3ml or 25ml overnight cultures using a QIAgen mini kit or QIAgen midi kit (Qiagen) respectively. Bacterial colonies or glycerol stocks were picked directly into bijoux's containing sterile 2xTY broth (Appendix I) supplemented with ampicillin at 100mg/ml. Cultures were grown overnight at 37°C in an orbital shaker and plasmid DNA was isolated using modified alkaline lysis kits according to the manufacturers instructions (Qiagen). Purified DNA was resuspended in sterile dH_20 and stored at -20°C.

2.2.3.10 Restriction endonuclease digestion verification of HIV-2 LTR clone construct

The presence of the HIV-2 LTR in the purified reporter construct was verified by restriction endonuclease digestion using the two endonucleases used to clone the insert into the construct. Digestion reactions were carried out in 10µl volumes for 2 hours at 37°C. 1µl of purified LTR construct: 0.5 Units *Mlu I*, 0.5 Units *Bgl II*, 1µl 10x Buffer H (Appendix 1) (Boehringer Manheim), was made to the total reaction volume with dH_2O. Following incubation the mixture was loaded onto a 1% agarose gel containing 0.5mg/ml ethidium bromide (Appendix I) and electrophoresed at 100v for 30 mins in 1xTAE buffer (Appendix I) to ensure fragment resolution.

2.2.4 Mammalian Cell Culture Techniques

2.2.4.1 Adherent cell lines

2.2.4.1a 293 Human Epithelial Kidney Cells

293 cells derived from human embryonic kidney cells were cultured in complete tissue culture medium DMEM (Appendix I), in T75 vented top flasks (Falcon) and incubated horizontally at 37°C in a 5% CO_2

humid atmosphere. Cells were split when approaching 90% confluence and maintained at a concentration of between $2x10^5$ and $2x10^6$ cells per ml. Monolayers were washed once with sterile PBS (Appendix I) after which the cells were overlaid with 2mls of 0.5μM filter sterile PBS 0.05% 0.5M EDTA. The monolayers were incubated under these conditions for 10 minutes at 37°C following which cells were removed from the surface of the plastic by several sharp smacks of the flask against the palm. The EDTA containing cell mixture was then pipetted up and down to dissociate cell clumps and an aliquot was removed and placed into a clean flask containing 15mls of complete media (Appendix I).

2.2.4.2 Suspension cell lines

2.2.4.2a Jurkat and THP-1

Jurkat, a T cell line derived from Human T-cell leukaemia, and THP-1, a monocyte like cell line derived from Human acute monocytic leukaemia, were grown in T75 and T150 vented top flasks (Falcon) containing 30ml or 80ml complete media RPMI (Appendix I) respectively. Flasks were incubated at 37°C in CO_2 humid atmosphere and were passaged approximately every three days. Cells were pelleted by centrifugation at 1200rpm for 8 minutes in sterile 20ml tubes, pellets were washed once in warm media and cells were then re-pelleted. Finally, pellets were resuspended in 1 ml of complete media and aliquots of cell suspension were transferred into clean flasks containing complete media RPMI (Appendix I).

2.2.4.3 Cryogenic storage of mammalian cell lines

Aliquots of adherent and suspension cell lines were preserved long term by storage in liquid nitrogen. Cells were pelleted by centrifugation at 1200rpm for 10 minutes in sterile 20 ml tubes. Cell pellets were resuspended in 1ml of freezing media (10% DMSO, 40% FCS, 50% RPMI or DMEM) and transferred to Nunc Cryovials (Gibco BRL). Cells were placed at -70°C overnight after which they were transferred to a liquid nitrogen stack for long term storage.

2.2.4.4 Transfection of adherent cell lines

2.2.4.4a Calcium phosphate transfection

24 hours prior to transfection adherent cell monolayers at 70% confluence 12 well dishes were seeded at a concentration of approximately 1 x 10^5 cells per ml. After 24 hours, media was removed from each well and cell monolayers were washed twice with filter sterile PBS (Appendix I), overlaid with fresh serum free medium and incubated for 3-4 hours at 37°C. Each monolayer was transfected with 2.5μg of

DNA consisting of "experimental" reporters or substituted with Herring Sperm DNA (HSDNA). DNA was mixed with 6.2µl 2M $CaCl_2$ and distilled water to a total volume of 50µl in a sterile plastic bijoux. This mixture was combined with 50µl of 2 x Hepes Buffered Saline (Appendix I) (added dropwise), complete incorporation was ensured by rapid air bubbling of the mixture with a pipette. Incubation at room temperature for 20 minutes allowed the formation of 100µl of white precipitate that was added dropwise onto the cell monolayers. Plates were incubated at 37°C for a further 4 hours, after which time the medium was removed and monolayers were washed once with one volume of filter sterile PBS (Appendix I) followed by one volume of serum free medium. Cell monolayers were overlaid with 1 ml of serum free DMEM and incubated for 12-16 hours at 37°C, this medium was replaced with 1 ml of complete 10% FCS DMEM (Appendix I) and cells were incubated for a further 24 hours. Cell monolayers were harvested 36-48 hours post transfection as described in section 2.2.4.5.

2.2.4.5 Harvesting of transfected adherent cell lines

Prior to harvesting adherent cell monolayers, medium was removed from each well and monolayers washed twice, gently, with one volume of filter sterile PBS (Appendix I). Cells were overlaid with 250µl of Passive Lysis Buffer (Promega Ltd.) and placed on a rocking table for 30 minutes. Cell lysates were pipetted up and down to disperse any remaining attached cells and transferred to 0.5ml polypropylene tubes. Cell debris was pelleted by pulse centrifugation and lysates were stored at -70°C.

2.2.4.6 Transfection of suspension cell lines

2.2.4.6a SuperFect™ Transfection

24 hours prior to transfection growing suspension cells at concentrations of between $2x10^5$ and $2x10^6$ cells per ml were passaged into fresh T75/T150 cm^2 flasks at a 1:2 ratio as described in section 2.2.4.2a. On the day of transfection cells were pelleted, resuspended and seeded into 6 well dishes, each well receiving 2 x 10^6 cells in 1.6 ml of complete 10% FCS RPMI (Appendix I), and incubated at 37°C and 5% CO_2. 2µg of DNA dissolved in sterile distilled water were diluted with RPMI containing no serum, proteins or antibiotics to a total volume of 100µl in a sterile eppendorf. This solution was vortexed for 3 seconds to ensure uniform mixing. 8µl of SuperFect™ (Qiagen) was added to the DNA solution that was subsequently vortexed for 10 seconds and incubated at room temperature for 5-10 minutes allowing the activated dendrimer - DNA complexes to form. Following incubation, 1ml of complete 10% FCS RPMI was added to the tube containing the transfection complexes and the mixture was pipetted up and down

twice, then immediately added dropwise onto the cells. Uniform distribution of the complexes in the well was achieved by gently swirling the dish. Plates were incubated at 37°C and 5% CO_2. 36-48 hours post transfection cells were harvested by centrifugation as described in section 2.2.4.7.

2.2.4.6b Effectene™ Transfection

Cells were prepared as described above. 2µg of DNA dissolved in sterile distilled water were diluted with DNA-condensation buffer EC, to a total volume of 100µl in a sterile eppendorf containing 3.2µl of Enhancer. This solution was vortexed for 3 seconds to ensure uniform mixing. 4µl of Effectene™ were added to the DNA solution that was subsequently vortexed for 10 seconds and incubated at room temperature for 5-10 minutes allowing non-liposomal lipid - DNA complexes to form. Following incubation, 1ml of complete 10% FCS RPMI was added to the tube containing the transfection complexes and the mixture was pipetted up and down twice, then immediately added dropwise onto the cells. Uniform distribution of the complexes in the well was achieved by gently swirling the dish. Plates were incubated at 37°C and 5% CO_2. Post transfection procedures were as described in section 2.2.4.7.

2.2.4.7 Harvesting of transfected suspension cell lines

36-48 hours post transfection plates cells were harvested. Using a plastic Pasteur pipette cells were transferred from the well into a plastic 15ml conical centrifuge tube. Cells were pelleted by centrifugation at 1200rpm for 7 minutes where after medium was removed and the pellet was resuspended with 100µl of Passive Lysis Buffer (Promega Ltd.). Cell lysates were transferred to 0.5ml polypropylene tubes and cleared of cell debris by pulse centrifugation, lysates were stored at -70°C.

2.2.5 Detection of reporter gene expression

2.2.5.1 Determination of LTR activity by Dual Luciferase Assay

The functional activity of each LTR clone was taken as a measurement of firefly luciferase expression. Transfection efficiency was assessed by *Renilla* luciferase expression and used to normalise LTR activity values. By measuring chemiluminescence a dual luciferase assay was used to quantify the expression from each of the two luciferase reporter vectors within prepared cell lysates. Firefly luciferase has an enzyme structure distinct from *Renilla* luciferase and has, therefore, different substrate requirements. The photon signal produced from firefly luciferase was achieved upon addition of luciferin and ATP. This reaction was quenched while simultaneously activating photon emission from *Renilla* luciferase by a

reagent containing coelenterazine. The assay was performed using 20µl of thawed lysate dispensed into the well of a black 96 well flat-bottomed plate (Nunc). The firefly luciferase reporter assay was initiated by the addition of 100µl of Luciferase Assay Reagent II (LARII) (Promega Ltd) to the well (Figure 2.2, pg 122). Chemiluminescence produced from the firefly luciferase reaction was measured immediately using a 96 well plate reading Canberra-Packard Luminometer. Quenching of firefly luciferase luminescence, and concomitant activation of *Renilla* luciferase was accomplished by the addition of 100µl of Stop & Glo™ Reagent (Promega Ltd) to each well (Figure 2.2, pg 122). The luminescent signal produced from the *Renilla* luciferase reaction was measured immediately also using a 96 well plate reading Canberra-Packard Luminometer. The luminescent signal from the firefly reaction was quenched by at least a factor of 10^5 (to $\leq$ 0.001% residual light output) within 1 second following the addition of Stop & Glo™ Reagent. Complete activation of *Renilla* luciferase is also achieved within the 1-second period. LTR activity is represented by the firefly luminescent signal divided or "normalised" by the *Renilla* luminescent signal. The assay could be performed immediately following lysate preparation or following the storage of the lysate at -20°C.

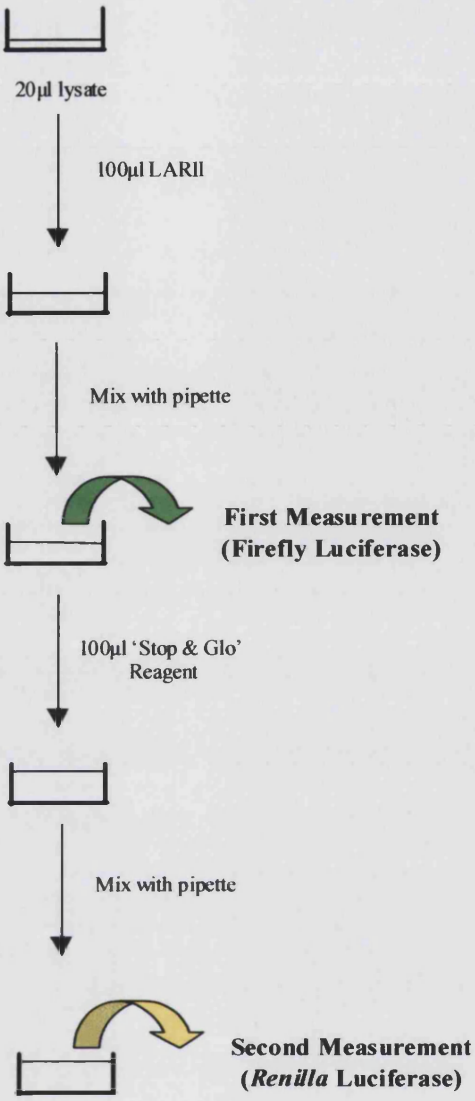


Figure 2.2 Schematic diagram of the dual luciferase assay. Reactions were performed in a black 96 well flat-bottomed plate (Nunc) to avoid cross-talk or leaching of luminescence between wells. LARII and Stop & Glo™ reagents were added using a multichannel pipette. Each plate reading was completed in approximately 5 minutes.

2.2.5.2 Determination of background firefly and Renilla luciferase activity by Dual Luciferase Assay

Background activity was assessed by assaying the cell lysate from a transfection of firefly luciferase reporter vector pGL3E lacking an HIV-2 LTR, made up to a total DNA concentration (according to the cell-line used) with HSDNA (an internal control plasmid lacking a promoter region was not available as negative control DNA for the pRL-TK plasmid). Two background activity measurements were made. Firstly, firefly luciferase background was measured following the addition of the firefly luciferase substrate (DLAI) to the cell lysate; secondly, *Renilla* luciferase background was measured following the addition of the *Renilla* luciferase substrate (DLAII). The first background measurement was subtracted from all firefly luciferase measurements; the second background measurement was subtracted from all *Renilla* luciferase measurements.

2.2.6 Automated DNA dye-primer cycle sequencing

Sequencing of all HIV-2 LTR clones was undertaken utilizing the Thermo Sequenase™ pre-mixed cycle sequencing kit (Amersham International plc) for use with the ™Vistra DNA Sequencer 725 (Amersham International plc, Molecular Dynamics Inc, USA). The sequencing protocol involved repeated cycles of thermal denaturation, annealing, and extension/termination reactions. Fluorescent dye-primers were used in the sequencing reactions in conjunction with dideoxynucleotides (ddNTP's); ddNTP incorporation into newly synthesised DNA resulted in chain termination. Electrophoresis of reaction products was carried out on a denaturing polyacrylamide/urea gel; fluorescence from the labelled DNA was detected using the ™Vistra DNA Sequencer 725. High quality template DNA was essential to obtain the best results from the kit, and was prepared using the Qiagen Midi kit (Qiagen) as described in section 2.2.3.9. The sequencing protocol used resulted in ≥ 98.5% accurately read bases, 400bp beyond the 3'-end of the primer used.

2.2.6.1 Sequencing Oligonucleotides

Dideoxynucleotide sequencing reactions were performed using Texas Red™ (Oswel) labelled primers that had a single fluorescent moiety covalently coupled to the 5' end. Texas Red™ is a fluorescent reporter molecule, which has an absorption maximum at 596nm and an emission maximum at 615nm. The incorporation of a Texas Red™ coupled primer into newly synthesised DNA results in the ability to detect reaction product sequences by fluorescence, a quicker and safer alternative to using radioactivity. The sequencing protocol routinely resulted in 400 accurately sequenced bases from the 3' end of the primer, thus to ensure the entire 758bp cloned LTR region was sequenced, three fluorescent dye labelled primers were designed. Two primers were designed to flank the cloned LTR region and were therefore located within the pGL3E firefly luciferase vector. The third primer was designed to bind within the cloned LTR region; this conferred overlap of the flanking sequencing fragments and enabled a complete and accurate sequence to be obtained. All primers used in the sequencing reactions are listed in Table 2.6, pg 124.

Primers for DNA sequencing of HIV-2 Long Terminal Repeat (LTR)

All primers used in the sequencing protocol described in section 2.2.6 were labelled with Texas Red™ dye and were synthesised by Oswel.

Primer Name	Orientation	Location	Sequence
C2280	Sense	5006-5025[a]	5' CTAGCAAAATAGGCTGTCCC 3'
C2281	Antisense	89-111[a]	5' CTTTATGTTTTTGGCGTCTTCC 3'
Y5152	Sense	9244-9269[b]	5' CCAGAGGAATTTGGGCACAAGTCAGG 3'

Table 2.6 Oligonucleotide primer sequences for the cycle sequencing of HIV-2 LTR. [a]Sequences numbered according to the plasmid map for pGL3E, (Promega Ltd).
[b] Sequences numbered according to HIV-2 ROD, GenBank accession number M15390.

2.2.6.2 Sequencing of double stranded DNA

Two pmoles or six microlitres of the DNA template were added to 2pmoles (1 pmole/µl) of the appropriate sequencing primer and 18 µl of distilled water. This mixture was divided equally between four polypropylene microfuge tubes, each tube pre-dispensed with 2 µl of pre-mixed enzyme/termination reagent which consisted of either ddATP, ddGTP, ddCTP, or ddTTP and Thermo Sequenase DNA polymerase. When all of the reactions had been prepared they were overlaid with light mineral oil and pulse centrifuged for 10 seconds to ensure all reagents were below the mineral oil layer. The sequencing reactions were carried out using a Perkin Elmer thermal cycling machine. The thermocycling programme for the sequencing reactions was preceded by a pre-melting step of 94°C for 4 minutes. The conditions thereafter were 94°C for 30 seconds, 50 °C for 15 seconds and 72 °C for 30 seconds over 25 cycles followed by a 4 °C soak. Before loading reaction products on the gel, samples were mixed with loading dye (Appendix I) and the total volume of the reaction was reduced. Sequencing reactions were recovered from under the mineral oil and dispensed into the wells of a conical bottomed 96 well plate (Nunc), each well-contained 3µl of loading dye (Appendix I). The mixture was pipetted up and down to ensure uniform dispersal and the 96 well plate was placed on top of a pre-heated 80°C heat block. The plate and heat block were put into a vacuum desiccator until the reaction/loading dye mixture had been reduced from a volume of 11µl to a volume of 3µl. Reaction products were loaded onto a gel immediately or stored at 20 °C until required.

2.2.6.3 Sequence electrophoresis

The sequencing reaction products were resolved on a 6% denaturing polyacrylamide/urea gel (Appendix I). The glass sequencing plates were cleaned by washing first with an alkaline detergent solution of 10% Decon followed by distilled water and finally 70% ethanol, plates were dried with soft paper tissues. The plates were aligned according to their bevelled edges and spaced by one opaque plastic spacer and one transparent acrylic spacer with protruding tabs. The spacer with protruding tabs was positioned to align with the entry point of the ^{TM}Vistra DNA Sequencer 725 laser. The plates were placed in a horizontal position for gel pouring and were held together with bulldog clips.

The 6% denaturing polyacrylamide/urea gel was degassed for three minutes in a vacuum dessicator prior to the addition of 150µl of ammonium peroxodisulphate (Appendix I) and 15µl of TEMED. On addition of these two reagents the gel was poured immediately using a 50ml syringe, a plastic surface former was inserted between the gel plates and the gel was left to polymerise for 60 minutes. Once polymerised the

gel was placed in the ™Vistra DNA Sequencer 725 and the anode and cathode reservoirs were filled with 1x TBE buffer (Appendix I). The surface former was removed and the gel surface was flushed with reservoir buffer to remove excess urea.

Before sample loading, the gel was preheated by running it at 1000 volts and 17 milliamps for 10 minutes. Following this period a paper sharks tooth comb was inserted into the gel surface and 3µl aliquots of sequenced products were loaded into each well. All empty wells were aliquoted with 3µl of 2:1:1 loading dye/salt solution/water to prevent track distortion. The gel was run at 1400 volts at 22-27 milliamps for 9 hours in 1x TBE buffer (Appendix I).

2.2.6.4 Sequence detection

The ™Vistra DNA Sequencer 725 contained a low power, helium-neon laser that generated light at 594.1nm in the yellow region of the visible spectrum. At a distance of 30cm from the sample wells the DNA fragments from the sequencing reactions were illuminated by the laser beam, which entered through the protruding tabs of the transparent spacer and traversed the entire width of the gel. Flourescence from the labelled DNA passed through a window behind the gel and was focussed through a band-pass filter onto a 512-element photodiode array detector. The detected signal was digitalised and transmitted to the computer, which generated a fluorescent intensity profile of the gel. In this way the computer identified the position of each DNA band and automatically called the bases. The automatic base calling was checked manually and corrected where necessary. Overlapping sequence profiles were aligned into contiguous sequences by downloading the computer generated information into a sequencing programme called Sequencher (Gene Codes Corporation).

Chapter 3.

PCR Amplification of HIV-2 LTR genotypes.

3.1 Introduction

In order to investigate comprehensively the relationship between HIV-2 LTR function, the replicative phenotype of the virus and/or the progression status of the infected patient, it was necessary to produce accurate DNA copies of the LTR from virus originating from a variety sources that included phenotypically and clinically characterised viral isolates and patient-derived material. It is possible to amplify selectively specific DNA sequences using the Polymerase Chain Reaction (PCR) (Mullis & Faloona, 1987). This application involves a series of denaturation, annealing, and extension reactions primed by specific oligonucleotide sequences. The following chapter describes how this technique has been used to target and amplify LTR sequences from molecular clones, viral isolates, and patient-derived material.

The clinical material of interest in this study was derived from patients with disparate proviral DNA loads relating to their disease progression rates (see section 3.2.1.3). Proviral load measurements from asymptomatic non-progressing patients were extremely low (<1000 copies per 10^5 $CD4^+$ cells), whereas rapidly progressing patients exhibited proviral loads of >50,000 copies per 10^5 $CD4^+$ cells (determined by Dr. N. Berry, NIBSC, see section 2.2.1.3). PCR amplification from DNA containing high numbers of proviral copies did not prove to be problematic, however amplification of the LTR region from LTNP patients was much less successful. Optimisation of first and second round PCR conditions was required in order to obtain LTR amplicons from low copy number LTNP PBMC material. This chapter describes the optimisation and application of a nested PCR that is capable of amplifying the HIV-2 LTR from a single copy of proviral DNA using molecular clone, viral isolate, or patient PBMC derived DNA as starting material.

3.2 Establishment of a nested PCR for the amplification of the HIV–2 Long Terminal Repeat (LTR)

3.2.1 HIV-2 sources for Long Terminal Repeat PCR amplification

The HIV-2 LTR region was amplified from a wide variety of sources including, molecular clone (n=1), viral isolate (n=5), and patient material (n=6). Molecular clone and viral isolates were characterised phenotypically (Clavel *et al.* 1986b, Shultz *et al.,* 1990), and patient sources were characterised clinically (K.Ariyoshi & N.Berry, personal communication). This was essential in an evaluation of the relationship between LTR function and HIV-2 pathogenesis. The HIV-2 LTR was PCR amplified from the following sources:

3.2.1.1 HIV-2 Molecular Clone

$pROD_{10}$ - A full-length infectious clone, containing the complete proviral sequence from the isolate HIV-2 ROD (Clavel *et al*, 1986a, Clavel *et al*, 1986b).

3.2.1.2 HIV-2 Isolates

CBL Series - Chester Beatty Laboratory (CBL) 20-24, five viral isolates (MRC ADP) obtained from five different HIV-2 infected Gambian patients. The five Gambian HIV-2 strains (CBL 20-24) exhibit a range of *in vitro* phenotypes that have been correlated to the clinical status of the patients at the time of sampling. The phenotypes range from a highly cytopathic and fast replicating virus from a patient who died of AIDS (CBL 20), to non-cytopathic slow replicating virus from an asymptomatic patient (CBL 24) (Schulz *et al.,* 1990) (see Chapter 5 section 5.2 & Table 5.1, pg 192). Isolates were grown on H9 or C8166 cells before DNA extraction (performed by colleagues at the Chester Beatty Laboratories, see section 2.2.2.1a) and LTR cloning.

3.2.1.3 HIV-2 Clinical Material

The Gambian Cohort - PBMC samples were obtained from six HIV-2 seropositive patients attending the MRC hospital in Fajara, The Gambia. Each of the six patients had been recruited to a longitudinal study in which patient samples were taken at six monthly intervals (Ariyoshi *et al.*, 2000). Disease progression rates were determined based on clinical data, decline of CD4% over the period of follow-up (determined by colleagues at the MRC laboratories, Fajara, The Gambia, see section 2.2.1.1), and plasma viral load measurements (performed by Dr. N. Berry, NIBSC, London, see section 2.2.1.3). Rapid progressor (RP)

patients were defined as those with CD4 declining from greater than 20% to less than 13% within 20 months, and whose plasma viral load measurements exceeded 50,000 copies/ml (RP1-RP3). Long-term non-progressor (LTNP) patients had stable CD4 counts for more than 56 months, and had plasma viral load measurements of <1000 copies/ml (LTNP1-LTNP3) (see Table 3.1, pg 132, and Chapter 6 section 6.2). DNA was extracted from PBMC samples using a Qiagen™ Blood Kit (Qiagen) (see section 2.2.2.1b) and quantified by fluorometry (see section 2.2.2.2). LTRs were cloned from PBMC samples obtained as close to the time of RNA viral load measurements as possible.

3.2.2 Generation of HIV-2 Long Terminal Repeat PCR Products

Several publications have demonstrated that the functionally relevant transcription factor binding sites in the HIV-2 LTR are located within the U3 and R sections of the promoter (Guyader *et al* 1987, Arya & Gallo, 1988, Markovitz *et al.,* 1990, Arya, 1991, Markovitz *et al.,* 1992, Hannibal *et al.,* 1993, Hannibal *et al.,* 1994, Clark *et al.,* 1995). Oligonucleotide primers were designed to target a 764bp region of the HIV-2 LTR encompassing the U3 and R segments of the LTR in addition to 35bps of the U5 section of the promoter. HIV-2 amplicons spanned nucleotide positions –556 to +208 of the promoter relative to the start site of transcription (see Figure 1.6, pg 71). Primer sequences were determined following reference to published HIV-2 sequences recorded in the Los Alamos Database (see Table 2.5, pg 110). Theoretical melting temperatures (Tms) and DNA polymerase requirements were also considered in the design of the primers (see section 2.2.2.3a). All oligonucleotide primers were between 18 and 25 nucleotides in length and had an approximately 50% G/C content (Table 2.5, pg 110). Oligonucleotides were synthesised by Oswel DNA Service (University of Southampton). The Tm for each primer was calculated according to:

$$Tm = 4\ (G+C) + (A+T) \text{ minus } 5^{\circ}C$$

The second round cloning primers were designed to contain 5' restriction endonuclease sites, which matched sites found within the polylinker region of the reporter vector chosen for use in this study (see Appendix II Figure II.1, pg 348). The primers designed for use in this study are listed in Chapter 2 Table 2.5, pg 110.

3.2.2.1 PCR Amplification of the HIV-2 LTR from Viral Isolate and Molecular Clone material

Amplification from virus isolate and molecular clone material was attempted using DNA derived from each of the five CBL isolates (20-24) and the molecular clone $pROD_{10}$ (see section 2.2.2.3b). A product equivalent to the expected amplicon of 764bp was generated from each of the six DNA samples (Figure 3.1, pg 131). PCR products were extracted and gel purified for restriction endonuclease cloning according to the protocol set out in 2.2.2.6.

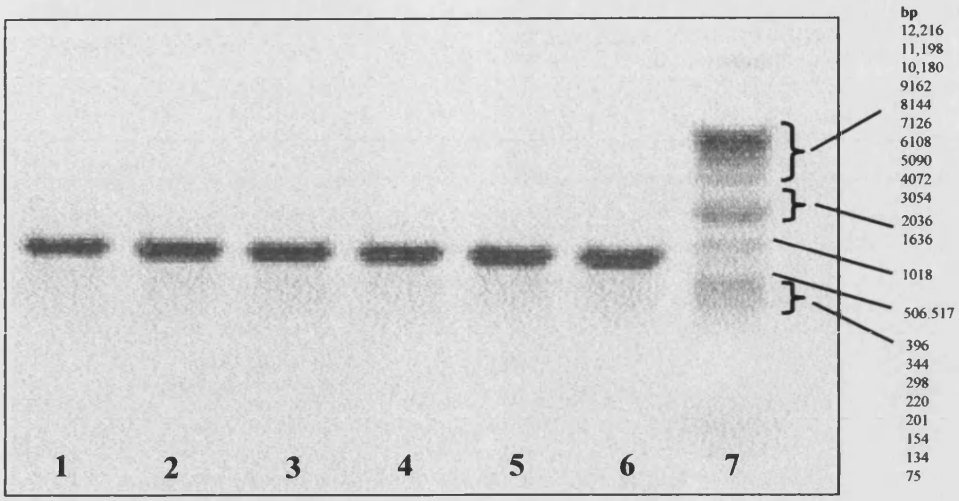


Figure 3.1 Nested PCR amplification of 764bp HIV-2 LTR from viral isolate and molecular clone sources using second round primers T5169 and T5170. Track 1: Molecular clone pROD10, 2-6 Viral isolate DNA samples (2: CBL 20, 3: CBL21, 4: CBL22, 5: CBL23, 6: CBL24), 7: DNA marker.

3.2.2.2 PCR amplification of the HIV-2 LTR from PBMC-derived patient DNA (Gambian cohort)

PCR amplification was attempted using PBMC-derived DNA from six different Gambian patients, three patients characterised as rapid progressors (RPs), the remaining three patients as long-term non-progressors (LTNPs) (see section 2.2.2.3b). HIV-2 LTR PCR products were generated from three of the six original samples, RP1, RP2, and RP3, the three patients characterised as RPs (Figure 3.2, pg 132). Partially successful PCR amplification was achieved with one LTNP sample (LTNP3), represented by a faint band seen in track 3 (Figure 3.2 – arrowed, pg 132), the remaining two LTNP samples (LTNP1 – track 9, and LTNP2 – track 4) failed to amplify under the conditions of the PCR protocol described in section 2.2.2.3b. Successfully amplified PCR products were extracted and gel purified for restriction endonuclease cloning according to the protocol set out in 2.2.2.6.

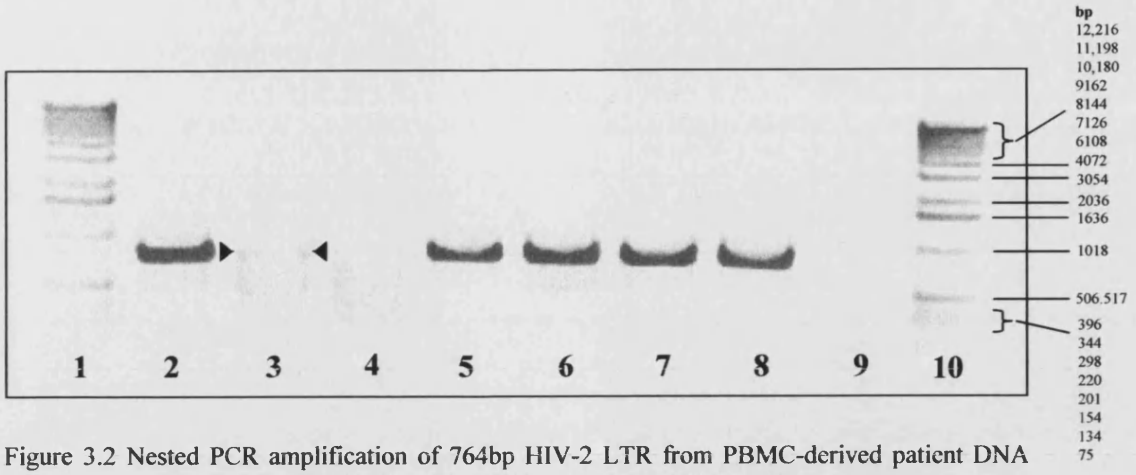


Figure 3.2 Nested PCR amplification of 764bp HIV-2 LTR from PBMC-derived patient DNA using second round primers T5169 and T5170. Tracks 1:DNA marker, 2:molecular clone pROD10 (positive control), 3-7, and 9 Gambian patient DNA (3:LTNP3, 4:LTNP2, 5:RP2, 6:RP1, 7:RP3, 9:LTNP1, 10:DNA marker).

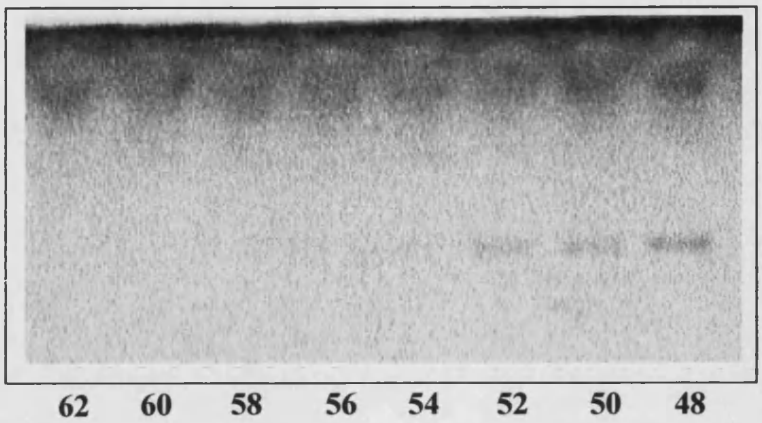
DNA Sample	CD4% (determined by colleagues at the MRC Laboratories, Fajara, The Gambia, see section 2.2.1.1)	Plasma HIV-2 RNA levels (copies per ml) (determined by Dr.N.Berry, see section 2.2.1.3)	Amplification of LTR from proviral DNA
RP1	**9**	**103000**	**Positive**
RP2	**22**	**55450**	**positive**
RP3	**13**	**183000**	**positive**
LTNP1	**41**	**undetectable**	**negative**
LTNP2	**49**	**undetectable**	**negative**
LTNP3	**48**	**undetectable**	**weak**

Table 3.1 Summary of clinical parameters of Gambian cohort patients. Undetectable RNA loads are considered to be <200copies/ml. Amplification success of the HIV-2 LTR region from patient proviral DNA is indicated.

3.2.2.3 Optimisation of the HIV-2 LTR nested PCR

In an attempt to increase the sensitivity of the PCR, modifications were made to the first and second round amplification reactions. PCR optimisation was performed using DNA derived from LTNP3 since partial amplification had been achieved with DNA from this patient. Initially changes to magnesium concentration and DNA polymerase were made, however, amplification was not improved (data not shown). Following this, the amount of patient template introduced into the first round reaction mixture was increased from 300ng to 600ng (quantified by fluorometry, section 2.2.2.2) and second round reactions were performed at a range of different annealing temperatures in order to identify the experimental optimum temperature for the primer pairing (see 2.2.2.4a). Second round PCR reactions had previously been performed at an annealing temperature of 54^{0}C. In this set of experiments second round reactions were performed at temperatures ranging from 48°C to 62°C, increasing by increments of 2°C with each reaction.

A. 300ng



B. 600ng

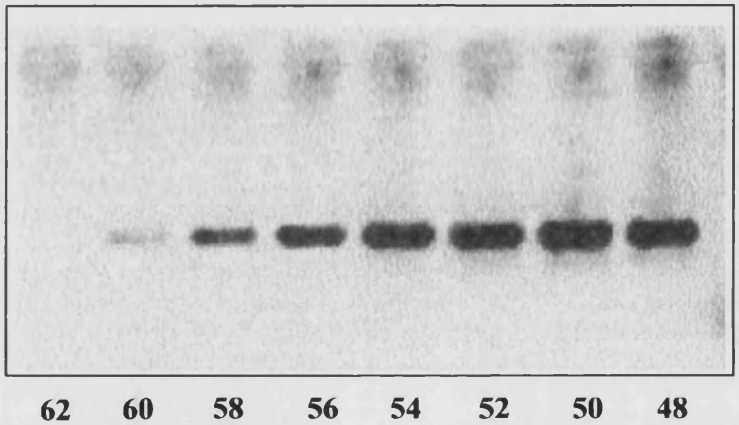


Figure 3.3 Nested PCR amplification performed using (A) 300ng and (B) 600ng PBMC-derived DNA from patient LTNP3. Reactions were performed over a second round annealing temperature gradient from 48°C to 62°C. (A) 300ng DNA PCR, Tracks 1-8: annealing temperatures 62°C to 48°C in increments of 2°C. (B) 600ng DNA PCR, Tracks 1-8: annealing temperatures 62°C to 48°C in increments of 2°C.

Amplification of the 764bp LTR region was found to be more successful when the first round DNA template input was 600ng as opposed to 300ng, regardless of the second round annealing temperature (see Figure 3.3, track 5 (A) vs. track 5 Figure 3.3 (B), track 6 (A) vs. track 6 (B), pg 134). However, differences in amplification between the annealing temperatures within the range were also demonstrated. PCR amplification of the 764bp LTR fragment was found to increase with each 2°C decrease in second round reaction temperature (Figure 3.3, pg 134). Optimal PCR amplification of the 764bp LTR fragment was achieved at a second round annealing temperature of 50°C (Track 7 Figure 3.3 (B), pg 134), with a 600ng DNA input.

3.2.2.4 Assessment of optimised HIV-2 LTR PCR performance by limiting dilution

Ultimately the goal of PCR optimisation is to achieve single copy sensitivity, that being the ability of the chosen PCR assay to produce detectable amplicons from a single HIV-2 molecule. A nested PCR capable of such sensitivity would increase the likelihood of successful amplification from low copy number material. In order to assess the single copy performance of the newly optimised PCR, limiting dilution analysis was performed (2.2.2.4b). A PCR that demonstrates a Poisson distribution of end-point molecules by limiting dilution, on the basis of probability represents single molecule amplification (Simmonds *et al.*, 1990).

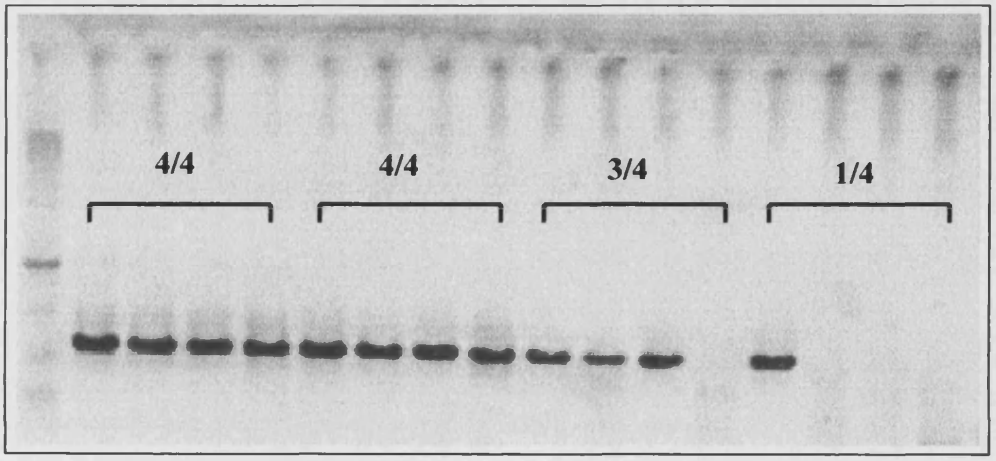


Figure 3.4 Limiting dilution PCR performed using PBMC-derived DNA from LTNP3 and second round primers T5169 and T5170. Tracks 1:DNA marker, 2-5:DNA input volume = n, 6-9:DNA input volume = n/5, 10-13:DNA input volume = n/25, 14-17:DNA input volume = n/625. The frequency of negatives or Fo = 0.75 is achieved at a 625 fold dilution of the original DNA sample. n = 2µl of DNA.

The results shown in Figure 3.4, pg 135, clearly demonstrate a Poisson distribution of end-point molecules (Simmonds *et al.*, 1990). Using the optimised PCR protocol, the HIV-2 LTR PCR second round primer pairing was sensitive enough to produce an LTR amplicon from a single copy of the HIV-2 genome.

3.2.2.5 Amplification of HIV-2 LTR from low copy number patient material (Gambian cohort) using a limiting dilution sensitive nested PCR

The PCR amplification of the HIV-2 LTR from LTNP low copy number PBMC-derived DNA was attempted using the newly optimised nested PCR protocol (see section 2.2.2.5). HIV-2 LTR PCR products were obtained for all three samples, LTNP1, LTNP2, and LTNP3 (Figure 3.5, pg 136). PCR products were extracted and gel purified for restriction endonuclease cloning.

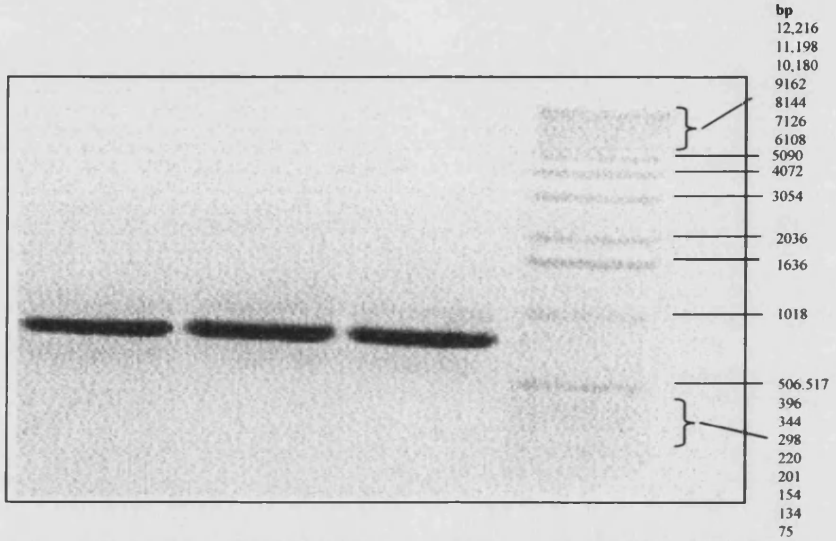


Figure 3.5 Limiting dilution sensitive nested PCR amplification of 764bp HIV-2 LTR from three LTNP Gambian patient DNA samples using second round primers T5169 and T5170. Tracks 1: LTNP1, 2: LTNP2, 3: LTNP3, 4: DNA marker.

3.3 Summary and discussion

The functional study of a genomic element necessarily requires the generation of an accurate and complete copy of the region in question. With the advent of PCR technology, generation of such a copy is possible through the direct amplification of a defined genomic sequence. In order to generate representative HIV-2 LTR regions, oligonucleotide primer sequences had to be designed to encompass defined LTR sequences. The parameters of the HIV-2 LTR have been determined in a number of studies based upon sequence and functional analysis (Guyader *et al.*, 1987, Arya & Gallo, 1988).

In 1987, Guyader *et al* published the genomic organisation and nucleotide sequence of an integrated proviral DNA from the ROD isolate of HIV-2 (Guyader *et al.*, 1987). The limits of the HIV-2 LTR and its internal U3, R, and U5 domains were defined by a combination of sequence analysis and complementary DNA synthesis. The U3 region was found be 556bp long, the R region 173 bp, and the U5 element 125bp. As stated previously, sequence elements of functional significance to gene expression within the HIV-2 LTR were identified in later studies (Arya & Gallo, 1988, Markovitz *et al.*, 1990, Tong-Starksen *et al.*, 1990, Arya, 1991, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995). The core enhancer region and upstream regulatory elements are contained within the 556bp of the U3 region (Guyader *et al.*, 1987, Arya & Gallo, 1988, Markovitz *et al.*, 1990, Tong-Starksen *et al.*, 1990, Arya, 1991, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995). The R region encodes a *trans*activation responsive element (TAR), which is essential for Tat *trans*activated gene expression (Guyader *et al* 1987, Arya & Gallo, 1988, Fenrick *et al.*, 1989, Chang & Jeang, 1992, Elangovan *et al.*, 1992, Rhim & Rice, 1993). However, mutational analysis of the 125bp making up the U5 region of the LTR has indicated that the sequences within this region do not significantly influence LTR directed gene expression (Arya, 1991).

For this study, oligonucleotide primers were designed to amplify the region of the HIV-2 LTR of functional significance to gene expression. The amplified LTR sequence represents 764bp of the original 854bp HIV-2 LTR, the amplicon contains the entire U3 (556bp) and R (173bp) regions of the LTR, in addition to 35bp of U5. The defined region of the HIV-2 LTR selected for functional analysis in this study is considered to be the functionally relevant transcriptional unit and these parameters have been the chosen for study in several other functional investigations (Guyader *et al.*, 1987, Arya & Gallo, 1988,

Markovitz *et al.*, 1990, Tong-Starksen *et al.*, 1990, Arya, 1991, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995).

PCR amplification of the HIV-2 LTR from molecular clone, viral isolate, and high copy number RP patient material was readily achieved using conditions of high nested PCR stringency. It is likely that the high copy number and good quality of the DNA templates derived from these sources resulted in successful PCR amplification. However, the initial PCR protocol failed to generate LTR PCR products from the DNA derived from the PBMCs of the three LTNP patients. Since extraction of LTNP and RP DNA was performed using the same materials and protocol (see section 2.2.2.1b), it is unlikely that PCR inhibitors from the extraction process caused the failure to amplify from LTNP-derived DNA. Moreover, carry over of inhibitors with an extract input of only 1-2µl into a 50µl first round PCR volume would be negligible.

Previous studies concerned with PCR amplification from HIV-2 infected PBMC material have experienced similar difficulties when trying to amplify from asymptomatic patient samples (Grankvist *et al.*, 1992, Simon *et al.*, 1993, Berry *et al.*, 1994). A study by Simon *et al* found that the ability to detect the HIV-2 genome by nested PCR was a reflection of the level of virus within the patient, and correlated with CD4 counts (Simon *et al.*, 1993). Positive PCR amplification was achieved with only one out of 14 patients with $CD4^+$ counts of >500 x 10^6/l, in contrast, amplification was achieved with six out of six patients with $CD4^+$ counts of <200 x 10^6/l. The inability to detect HIV-2 DNA by PCR from patients exhibiting high $CD4^+$ counts and no signs of disease progression in this and other studies has been attributed to the low numbers of the HIV-2 genome carried within the PBMCs of these infected individuals (Grankvist *et al.*, 1992, Simon *et al.*, 1993, Berry *et al.*, 1994).

A characteristic of the LTNP patients in this study (Table 3.1, pg 132), and in general (Simmonds *et al.*, 1990, Berry *et al.*, 1994, Ariyoshi *et al.*, 1996, Berry *et al.*, 1998, Popper *et al.*, 2000), is very low to undetectable proviral loads (1 copy/1000 - 1 copy/10,000 PBMCs); in contrast, rapidly progressing patients exhibit high proviral loads (up to 1 copy/10 PBMCs). It is likely therefore, that amplification from the three LTNP samples failed, in part, because of the low number of proviral copies of HIV-2 within the extracted PBMC-derived DNA used as first round template.

Nevertheless, the observation of partial PCR amplification seen in track 3 of Figure 3.2 (arrowed), pg 132, demonstrates that there was at least one copy of the HIV-2 template within the PCR mixture performed

with the DNA from LTNP3, and shows therefore that the conditions of the nested PCR described in section 2.2.2.3b were not optimal for the second round oligonucleotide primer pairing. In this situation the likelihood of successful PCR amplification is greatly reduced and it is probable that this was also a significant factor in the failure to amplify from the low copy number LTNP-derived material.

The aim of an optimal PCR is to achieve successful amplification from a single molecule of DNA (Simmonds *et al.*, 1990). To increase the likelihood of successful amplification from all low copy number LTNP samples, nested PCR conditions were further optimised in an attempt to achieve single copy sensitivity. Initially, changes were made to DNA polymerase and $MgCl_2$ concentrations. It was hoped that by increasing the concentrations of DNA polymerase and/or $MgCl_2$ within the PCR reaction mixture, the likelihood of successful PCR amplification would be increased since both components can be limiting in a PCR reaction. Amplification from the DNA template derived from LTNP3 was not achieved however, when these parameters were altered (data not shown), suggesting that free molecules of DNA polymerase and Mg^{2+} were present within the PCR mixture and that other factors were limiting the reaction.

Following $MgCl_2$ concentration, DNA polymerase, and buffer (especially salt), the DNA template concentration and PCR annealing temperatures exert a strong influence over the result of a PCR. These two parameters were optimised further in an attempt to improve the sensitivity of the nested PCR. First round template input was increased from 300ng to 600ng, and the second round annealing temperature was reduced by 4°C to 50 °C. These steps together effectively increased both the number of proviral targets and further enhanced the conditions for second round primer annealing. The data presented in Figure 3.3, pg 134, clearly demonstrates that PCR amplification of the HIV-2 LTR from low copy number template was greatly enhanced by the increasing initial template concentration and lowering the stringency of second round reaction conditions. The failure to amplify the HIV-2 LTR from LTNP PBMC-derived DNA under the initial nested PCR conditions likely resulted from the quantity of proviral DNA in the PBMC-derived DNA being too small, and the second round annealing temperature being too high for the second round primer pairing.

When performed using a limiting dilution method (Simmonds *et al.*, 1990), the optimised HIV-2 LTR nested PCR gave an end-point dilution series that demonstrated single copy sensitivity. All subsequent

PCR experiments were performed with 600ng of DNA template as a first round input, and second round annealing reactions were carried out at 50°C rather than 54°C.

Optimisation of the first and second round PCR conditions resulted in a nested PCR that is capable of amplifying the LTR region from a single molecule of HIV-2 derived from a wide variety of sources (Simmonds *et al.*, 1990). This enabled the successful PCR amplification of the HIV-2 LTR from molecular clone, isolate, and high and low copy number patient DNA.

Chapter 4.

Optimisation of Cloning, Transfection, and Reporter Gene Detection Procedures for the Measurement of HIV-2 LTR Activity.

4.1 Introduction

Characterisation of promoter elements can be successfully achieved through the use of the reporter assay. This technique involves attaching the promoter element under investigation to a gene encoding a reporter molecule; the amount of reporter protein synthesised under various conditions is presumed to reflect the ability of the inserted sequence to promote transcription. Reporter assays have been extensively used in the characterisation of retroviral promoter and enhancer elements, particularly in the analyses of both the HIV-1 (Golub *et al.*, 1990, Koken *et al.*, 1992, Estable *et al.*, 1996, Rousseau *et al.*, 1997 Zhang *et al.*, 1997b), and HIV-2 LTRs (Markovitz *et al.*, 1990, Tong-Starksen *et al.*, 1990, Arya, 1991, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995, Faulkner *et al.*, 2001). Initial studies utilised either β-galactosidase or Chloramphenicol Acetyltransferase (CAT) reporter assays, however, more recent analyses have involved the use of luciferase reporter systems that have greater sensitivity and a better dynamic range, and that are quicker to perform (Alam & Cook, 1990, Van Lint *et al.*, 1997, Zhang *et al.*, 1997, Jeeninga *et al.*, 2000).

The principal aim of this study was to assess and compare the activity of a variety of HIV-2 promoters of unknown activities; it was crucial therefore to select a rapid, sensitive, and reproducible reporter system capable of detecting a wide range of promoter activities. The advantages and disadvantages of the various reporter systems available are discussed in this chapter, and the features of the reporter assay system utilised in this study are outlined in the first section.

The procedure by which the genetic element under investigation is inserted into the reporter vector must be both rapid and reliable. The chapter describes the cloning of PCR amplified HIV-2 LTRs into the firefly luciferase reporter vector pGL3E in the second section.

The third section of this chapter describes the optimisation of a dual-luciferase reporter assay used to detect reporter gene expression in the measurement of HIV-2 LTR activity. Experiments to determine the linear range, stability, and reproducibility of the detection assay have been described.

An assay based upon a reporter system relies on the ability to introduce the promoter-containing reporter vector into the cell-type of interest. The process by which DNA is introduced into cells is called transfection. A transfection can be either transient or stable, and can be carried out by a number of different techniques. Transient transfections involve the introduction of DNA into mammalian cells followed by a 2-4 day incubation period in which transcription of the transfected reporter gene takes place. After the incubation period, transfected cells are harvested and can be analysed for the presence of the reporter molecule. Stable transfections result in the integration of the transfected gene within the chromosomal DNA, a process directed by antibiotic selection. The fourth and last section of this chapter discusses the vast array of techniques, reagents, and cell lines that are available for the transient transfection of DNA. Experiments to determine the most efficient and reproducible transfection protocol for each particular cell line used in the study are described.

4.2 Reporter assay systems for the measurement of promoter activity

There are a number of different reporter systems currently available, including β-galactosidase, Chloramphenicol Acetyltransferase (CAT) and luciferase reporter assays; all have particular advantages and disadvantages. Common to all systems, the promoter region under investigation must be placed directly upstream of the reporter gene, within a reporter vector lacking an endogenous promoter. Fundamental to the concept of the reporter assay, the reporter gene must direct transcription of a reporter protein that is absent from the host cell and does not alter host cell physiology. The assay utilised to detect the reporter molecule must be simple yet efficient and, above all, extremely sensitive. In addition, the assay must exhibit a broad linear range of detection enabling the analysis of both large and small changes in promoter activity.

4.2.1 Choice of mammalian cell line

Ideally, studies investigating the activity of the transcriptional regulatory elements within the retroviral genome would be carried out in the primary cells that the virus naturally infects. However, it is widely known that primary cells such as $CD4^+$ T lymphocytes and macrophages are refractory to many DNA transfection techniques. A small number of studies have successfully transfected primary cells, but at very low efficiencies (Horvat *et al.*, 1989, Hazan *et al.*, 1990). Moreover, successful transfection of HIV reporter vectors into primary T cells has been shown to require preactivation, confounding studies assessing transcriptional activity at the basal level and transcriptional responses to other T-cell activation signals.

The vast majority of reporter gene assays are performed by transfecting transiently reporter vectors into established mammalian cell lines. In general, these cell lines show higher transfection efficiencies than primary cells, and are widely used for transient transfection. These include HEK293, Jurkat, HL-60s, CEM, and THP-1 cell lines. When selecting a mammalian cell line for use in a reporter assay-based study two of the most important considerations are the transfection efficiency of the cell line and its biological relevance to the virus under investigation. The Jurkat and THP-1 cell lines have been extensively used in the transient transfection of reporter vectors containing HIV-1 and HIV-2 promoter regions (Markovitz *et al.*, 1992, Hilfinger *et al.*, 1993, Clark *et al.*, 1995, Estable *et al.*, 1996, Van Lint *et al.*, 1997). Both cell lines are non-adherent and show high transfection efficiencies. In addition, Jurkat is a T cell-like cell line and THP-1 is a monocyte-like cell line. Both primary T cells and monocytes are

infected by HIV-1 and HIV-2, and both types of cells have been shown to play important roles in HIV infection (Klatzmann *et al.*, 1984b, Ho *et al.*, 1986, Gendelman *et al.*, 1989, Ho *et al.*, 1995, Spina *et al.*, 1997, Woods *et al.*, 1997). The human embryonic kidney cell line (293HEK) has also been extensively used in transient transfection analysis due to its high transfection efficiency and tolerance of a wide range of transfection techniques. The HEK293 cell line has been chosen for its ease of use in several of the initial dual-luciferase assay optimisation experiments within this study, while the more biologically relevant Jurkat and THP-1 cell lines have been reserved for transfection experiments which assess and compare the activities of the PCR-amplified HIV-2 LTRs.

4.2.2 Choice of reporter assay system

The most attractive features of a luciferase based reporter assay are its sensitivity and speed. In comparison, the β-galactosidase (β-gal), and Chloramphenicol Acetyltransferase (CAT) reporter assays currently available are less sensitive and more time consuming (Alam & Cook, 1990), however, on average, they are cheaper. Most luciferase assays can measure enzyme expression over an extremely broad linear range that extends over at least eight orders of magnitude; by comparison, CAT and β-gal reporter assays display linear ranges of between four and six orders of magnitude. In a direct comparison with a CAT reporter assay Alam *et al* showed that the luciferase reporter assay was 100 fold more sensitive (Alam & Cook, 1990).

Another significant disadvantage displayed by both β-gal and CAT reporter assays is that they are end-point assays, consequently a prolonged period of incubation post-harvesting (12-16 hours) is required before reporter gene expression levels can be determined. In contrast, luciferase reporter assays do not require such incubation periods and therefore, are considerably quicker to perform, for example levels of luciferase gene expression can be measured within one hour post-harvesting. The luciferase assay does have two disadvantages, the first of these is related to cost, both the reagents and detection device i.e. a luminometer, required to perform the luciferase assay are more expensive than the corresponding materials needed to perform either the β-gal or CAT reporter assays. Secondly, although many manufacturers have improved the half-life of luciferase, the stability of the signal emitted by the reporter molecule is short lived. Potentially this creates detection problems, particularly if a large number of samples are being analysed in a single experiment.

Taking into account the advantages and disadvantages of all three assays the luciferase based reporter system was selected for use in this study. The superior sensitivity and speed of this assay are suited to this study since a large number of HIV-2 promoters of unknown activities were to be assessed. Choosing the luciferase system also avoided the use of radioactivity.

4.2.3 Features of the dual-luciferase reporter assay system

The luciferase based reporter assay system selected for use in this study involves the simultaneous cotransfection of two distinct luciferase reporter vectors, and is thus termed a dual luciferase assay. The first vector acts as an experimental reporter vector into which the promoter region under investigation is inserted; the second provides an internal control to which measurement of the experimental reporter is normalised. The process of normalisation effectively eliminates the inherent experimental variability that arises from differences in transfection efficiencies and the number and health of cultured cells. The experimental reporter vector in this system - pGL3E (Promega Ltd) (Appendix II, Figure II.1, pg 348) contains the firefly luciferase (*Photinus pyralis*) gene, which encodes a 61kDa monomeric enzyme that does not require post-translational processing and therefore functions as a genetic reporter immediately following translation. The second reporter vector encodes the sea-pansy or *Renilla* luciferase (*Renilla reniformis*) enzyme, a 36kDa monomeric protein distinct from firefly luciferase in evolutionary origin, structure, and substrate requirement. This makes it possible to discriminate selectively between the bioluminescent reactions catalysed by the two different luciferase enzymes, enabling both firefly and *Renilla* luciferase signals to be measured within a single well.

In order to determine the activity of a promoter using this system, the promoter under investigation is cloned into the firefly luciferase reporter vector (pGL3E) directly upstream of the firefly luciferase gene. Mammalian cells are then cotransfected with both the firefly and *Renilla* luciferase reporter vectors. Approximately 48 hours post-cotransfection the cells are harvested and assayed for the detection of both firefly and *Renilla* luciferase chemiluminescence in a single well assay performed in a 96 well plate (Figure 4.1, pg 147).

Firstly, the assay measures the chemiluminescent signal emitted from the reaction catalysed by the firefly luciferase enzyme encoded by the experimental vector pGL3E (Promega Ltd) following the addition of a beetle luciferin substrate. The firefly luciferase chemiluminescent signal is directly related to the level of firefly luciferase expression, which is in turn directly related to the activity of the inserted promoter under investigation. Secondly, the assay measures the chemiluminescent signal emitted from the reaction

catalysed by the *Renilla* luciferase enzyme encoded by the "internal control" reporter vector following the addition of a reagent that quenches firefly luciferase luminescence while simultaneously activating the luminescent reaction catalysed by *Renilla* luciferase. Each firefly luciferase reporter value is divided by the *Renilla* luciferase value generated by the cotransfected "internal control" reporter, giving a measure of the activity of the inserted promoter that is normalised for differences in transfection efficiency and the health and number of transfected cells.

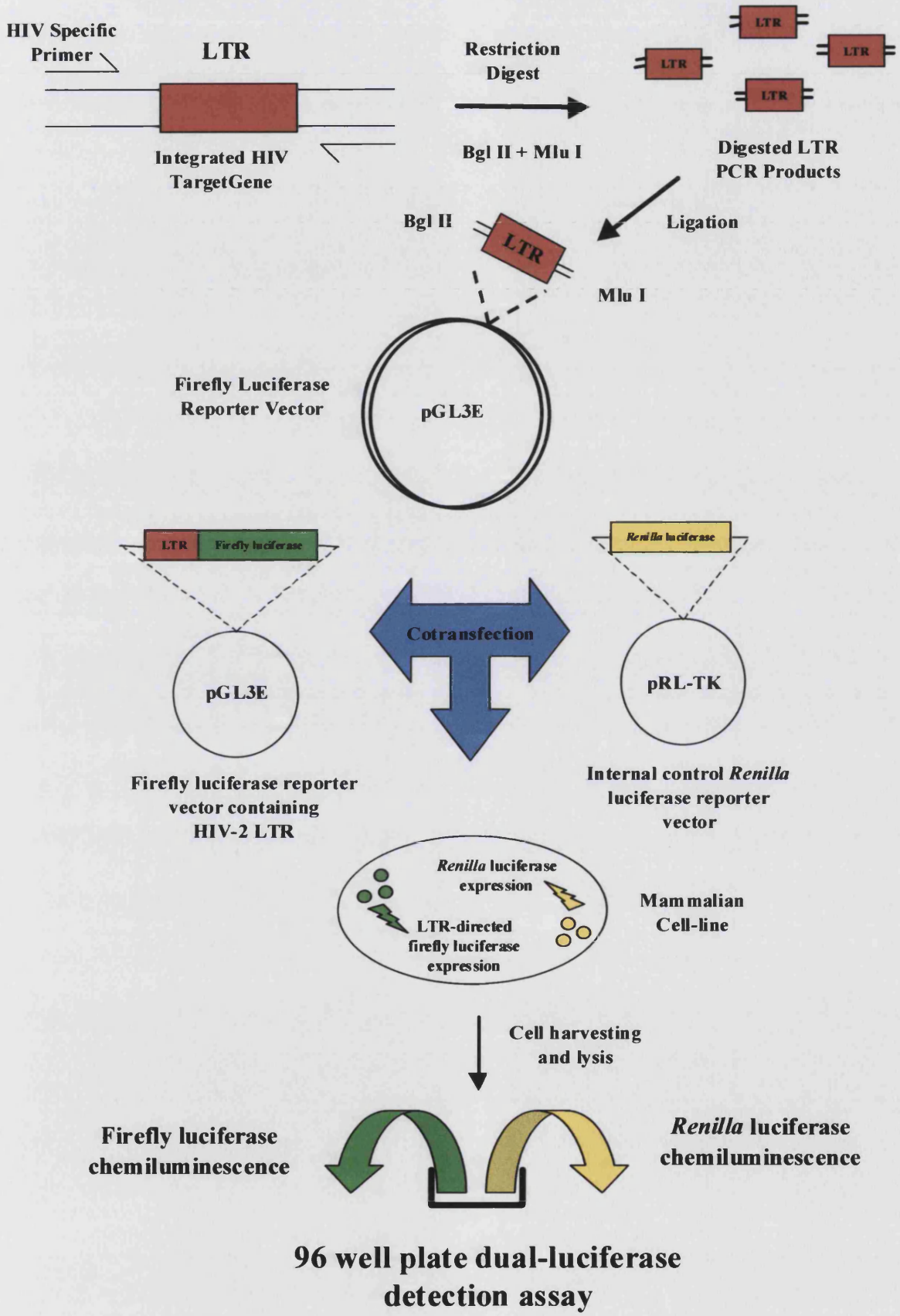


Figure 4.1 Schematic diagram outlining the procedures involved in the dual-luciferase reporter system used to measure HIV-2 LTR activity.

4.2.4 Features and requirements of the firefly luciferase reporter vector

Fundamental considerations when choosing a promoterless luciferase reporter vector included, luciferase expression levels directed by the vector, *in vivo* vector stability, levels of background luciferase expression directed by the vector, and flexibility to perform genetic manipulations. The firefly luciferase vector, pGL3-Enhancer (pGL3E, Promega Ltd) (Appendix II, II.1), was selected to assess quantitatively the activity of the LTR region derived from the various HIV-2 infected sources. The pGL3E reporter vector does lack a promoter region, but contains an SV40 enhancer located downstream of the firefly luciferase gene and the poly(A) signal. The inclusion of the SV40 enhancer is an important and unique feature of pGL3E, in that it acts to aid the verification of functional promoter elements inserted into the vector by increasing the overall level of firefly luciferase expression. The functional activities of the LTRs under investigation were unknown prior to the study, as was the transfectability of the cell lines to be used; therefore this particular feature of pGL3E increased the likelihood a signal being detected. A summary of the features which influenced the selection of pGL3E are written below:

- pGL3E contained an SV40 enhancer located downstream of the firefly luciferase gene and the poly(A) signal which aided the verification of functional promoter elements inserted into the vector by increasing the level of firefly luciferase gene transcription.
- pGL3E contained a multiple cloning site containing several restriction sites. This polylinker region was found upstream of the luciferase gene and allowed the rapid and efficient insertion of DNA promoter elements.
- An ampicillin-resistance gene, β-lactamase, and a high copy number prokaryotic origin of replication allowed pGL3E selection, maintenance, and manipulation in *E.coli* hosts.
- pGL3E contained the SV40 late poly(A) signal which increased the efficiency of transcription, termination and polyadenylation of the luciferase transcripts.

4.3 Cloning of HIV-2 LTR PCR products into the firefly luciferase reporter vector pGL3E

Amplified HIV-2 LTR PCR products were readily cloned into firefly luciferase reporter vector pGL3E using an efficient restriction endonuclease cloning protocol. Once optimised, this protocol achieved high frequencies of LTR positive clones from viral isolate, molecular clone, and patient derived amplification products.

4.3.1 Linearisation of pGL3E

To enable efficient cloning of PCR amplified LTRs into the firefly luciferase reporter vector it was first necessary to linearise the vector. This was achieved by restriction endonuclease digestion performed with two restriction endonucleases targeted to *MluI* and *BglII* sites within the polylinker region of pGL3E (Appendix II, Figure II.1, pg 348). The restriction endonuclease sites matched those designed into the second round PCR primers (Table 2.5, pg 110). Optimisation of the linearisation protocol was achieved by varying enzyme manufacturer, order of enzyme addition, and time of incubation. The efficiency of linearisation was evaluated by transforming 5µl of the resulting digestion mixture into *E.coli* and counting transformants.

The protocol resulting in optimal linearisation was; digest 1µg of pGL3E with 20 units *MluI* (Boehringer Mannhiem) in 1x Buffer H (Boehringer Mannhiem) for 1 hour at 37°C, residual enzyme was removed with 5µl Strataclean™ resin (Stratagene). This was followed by addition of 20 units *BglII* (Boehringer Mannhiem) and incubation at 37°C for a further 1 hour. Following restriction endonuclease digestion linearised pGL3E was dephosphorylated to prevent self-ligation. 1 Unit of Shrimp Alkaline Phosphatase (United States Biochemicals) was added to the digestion reaction mixture and incubated at 37°C for 20 mins. Inactivation of Phosphatase was achieved by a further incubation for 15 mins at 65°C. This protocol routinely gave rise to fewer than 8 transformants per microlitre of digestion product, indicating very efficient linearisation.

4.3.2 Restriction endonuclease digestion of amplified HIV-2 LTR PCR products

To enable the efficient cloning of PCR amplified LTRs into the firefly luciferase vector pGL3E, PCR products were also digested with restriction endonucleases *MluI* and *BglII* (Boehringer Mannhiem). The

PCR products contained sites for both enzymes since they were designed into the sequences of the second round PCR primers. The conditions for digestion of the PCR amplified LTRs are described in section 2.2.3.1.

4.3.3 Cloning of LTR PCR product into firefly luciferase reporter vector

Following digestion, the amplified LTRs were cloned into the linearised firefly reporter vector pGL3E; complementary annealing of the digested ends of both vector and PCR product resulted in the formation of a complete double-stranded circularised vector containing an LTR insert (Figure 4.1, pg 147). Optimisation of the ligation protocol was achieved by varying the vector to LTR insert ratio. Reaction ratios of 1:3, 1:5, and 1:10 (vector to insert) were routinely performed according to the protocol described in section 2.2.3.4. This range of ratios compensated for any variation in template concentration that arose from amplification or gel purification steps relating to the PCR product, or bacterial DNA purification techniques relating to the vector. The efficiency of ligation was evaluated by transforming 5µl of the resulting ligation mixture into *E.coli* and growing transformants on selective agar at 37°C overnight (see section 2.2.3.5). Colonies containing circularised vector possessing an LTR insert were screened by PCR (see section 2.2.3.7). The effects of varying the amount of vector to insert in the cloning reaction are shown in Figure 4.2, pg 151.

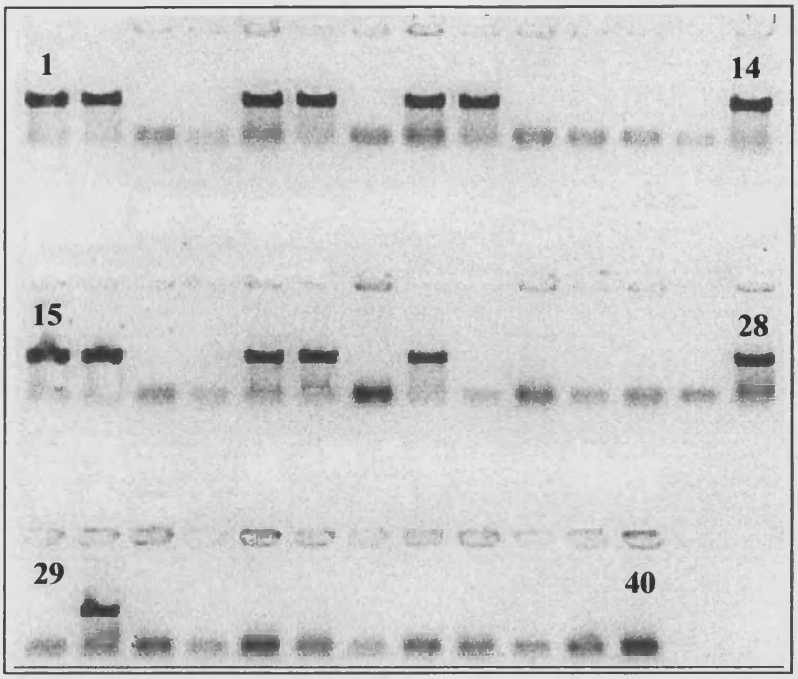


Figure 4.2 Gel photograph of PCR screen of bacterial transformants for HIV-2 LTR inserts. A 764bp band identifies LTR positive clones. Individual pGL3E vector batches routinely achieved cloning efficiencies of up to 80%. Tracks 1-20: (5µl transformation of 1:10 vector to insert cloning ratio, 11 positive colonies), 21-30: (5µl transformation of 1:5 vector to insert cloning ratio, 3 positive colonies), 31-40: (5µl transformation of 1:3 vector to insert cloning ratio, 0 positive colonies). DNA markers track is not shown.

These results clearly demonstrate that the efficiency of cloning was dramatically increased as the ratio of insert to vector was increased. In general, vector to insert ratios of 1:10 produced the highest frequency of HIV-2 LTR PCR positive colonies. This protocol routinely achieved cloning efficiencies of 70-80%.

4.3.4 Verification of LTR clone constructs by restriction endonuclease digestion

PCR positive colonies were picked into selective broth and cultured overnight at 37°C (see section 2.2.3.8), vector DNA was extracted from the overnight cultures using a Qiagen Midi Kit (Qiagen) (see section 2.2.3.9). Restriction endonuclease digestion reactions were performed with *BglII* and *MluI* (Boehringer Mannhiem) according to the protocol described in section 2.2.3.10, in order to verify the presence of the LTR insert within the reporter construct purified from overnight cultures. Following digestion, LTR containing constructs released a 764bp LTR fragment and a 5.064Kb vector backbone. The fragments released on digestion were resolved by agarose gel electrophoresis and appeared as two distinct bands correlating to the expected molecular weights (Figure 4.3, pg 152). All LTR clone constructs were verified in this way before being utilised in subsequent experiments. The number of clones obtained from each of the different sources is listed in Table 4.1, pg 152. Purified construct DNA was stored at -20°C until required.

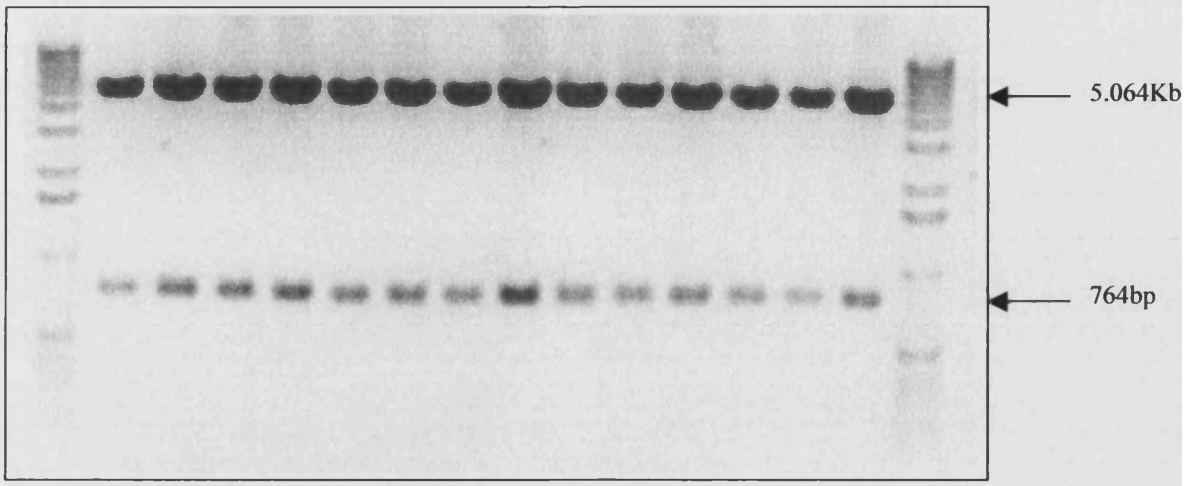


Figure 4.3 *BglII* and *MluI* digestion of pGL3E-LTR clones derived from the Gambian cohort. Tracks 1 and 16 DNA marker, 2-5: RP3 derived clones, 6-8: LTNP1 derived clones, 9-11: LTNP3 derived clones, 12-15: RP1 derived clones.

DNA Source	Total number of pGL3E-LTR clones
$pROD_{10}$	3
CBL20	7
CBL21	5
CBL22	10
CBL23	7
CBL24	13
LTNP1 (Gambian cohort)	8
LTNP2 (Gambian cohort)	9
LTNP3 (Gambian cohort)	8
RP1 (Gambian cohort)	10
RP2 (Gambian cohort)	8
RP3 (Gambian cohort)	7

Table 4.1 Summary of the number of pGL3E-HIV-2 LTR clones obtained from various HIV-2 sources

4.4 Evaluation and optimisation of the dual-luciferase detection assay

Before performing the large-scale cotransfection studies to assess and compare the activities of the PCR-amplified HIV-2 LTRs, several experiments were carried out to evaluate and optimise the dual-luciferase detection assay. Due to its high transfectability and rapid growth the HEK293 cell line was initially used to assess the optimal conditions for the detection of both firefly and *Renilla* luciferase. HEK293 cells were transfected in 12 well dishes according to the standard calcium phosphate transient transfection protocol described in section 2.2.4.4a. Subsequent sections of this chapter go on to describe the optimisation of Jurkat and THP-1 cotransfection procedures using the optimised dual-luciferase detection assay.

4.4.1 Stability and dynamic range of the firefly luciferase signal

Initially an experiment was carried out to determine the range of quantum efficiency and stability of the chemiluminescent signal emitted from the reaction catalysed by firefly luciferase. A serial dilution was carried out from a firefly luciferase stock concentration of 1μg/μl, such that chemiluminescence was measured from concentrations of firefly luciferase between 0.1pg and 10μg per well. Dilutions were made in Passive Lysis Buffer (Promega Ltd). 100μl aliquots of beetle luciferin firefly luciferase substrate (Promega Ltd) were added to each dilution and chemiluminescence was measured using a Canberra Packard scintillation counter. Following the initial measurement at one minute post-mixing, readings for each dilution were taken at five-minute intervals over the course of a twenty-minute period (Figure 4.4, pg 154). The results indicated that firefly luciferase chemiluminescence could be successfully detected over a linear range extending at least seven orders of magnitude. The data also suggested that the chemiluminescent signal was relatively stable over the entire linear range; only a two-fold drop in activity was observed over the twenty-minute period. The largest drop in activity over time was observed at the highest concentration of firefly luciferase, perhaps due to limiting amounts of substrate. However, since the Canberra Packard scintillation counter takes less than five minutes to read an entire 96 well plate, the results indicated that the drop in activity experienced by experimental samples in such a plate while measuring the firefly and *Renilla* luciferase signals would be negligible.

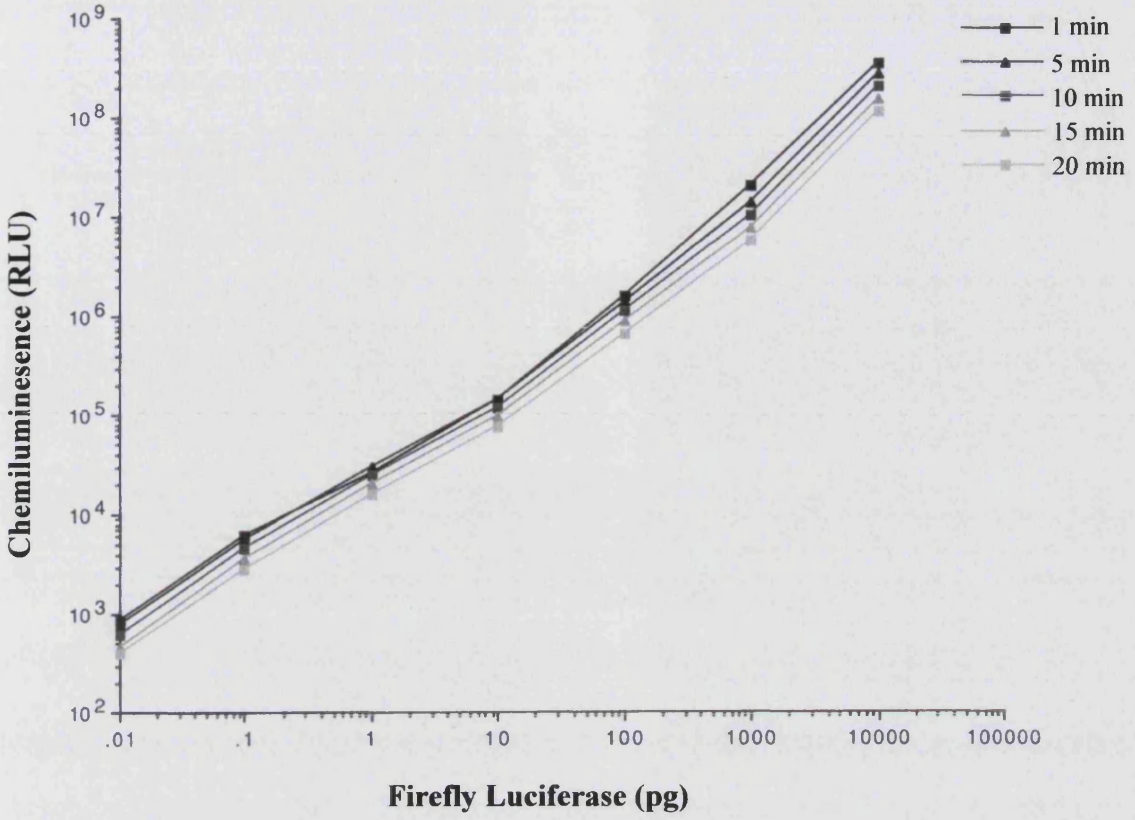


Fig 4.4 Titration curve of firefly luciferase chemiluminescence over time to test the dynamic range and stability of firefly luciferase detection using the dual luciferase assay system.

4.4.2 Efficiency of firefly luciferase quenching

The dual-luciferase detection assay is performed within a single well. The assay involves first detecting the firefly luciferase signal following the addition of a beetle luciferin substrate and is followed by the addition of a reagent termed 'Stop and Glo'™ (Promega Ltd) (see section 2.2.5.1), a chemical which simultaneously quenches the firefly luciferase-catalysed reaction while activating the *Renilla* luciferase-catalysed reaction, the chemiluminescence of which is then detected. In order to obtain an accurate *Renilla* luciferase reading the firefly luciferase reaction must be completely quenched by 'Stop and Glo'™. Any residual light output from the firefly luciferase reaction would artificially enhance the light output obtained from the *Renilla* luciferase reaction, the value of which is used to normalise the firefly luciferase value. It was therefore necessary to demonstrate that the firefly luciferase chemiluminescence was quenched effectively under the reaction conditions utilised in these studies. A 1:1000 dilution from a firefly luciferase stock concentration of 1μg/μl was made and added to a well of a 96 well plate. 100μl of beetle luciferin substrate was added to this sample and firefly luciferase chemiluminescence was measured. The reaction was then quenched by the addition of 100μl of 'Stop and Glo'™ reagent, and residual light output was measured (Figure 4.5, pg 155).

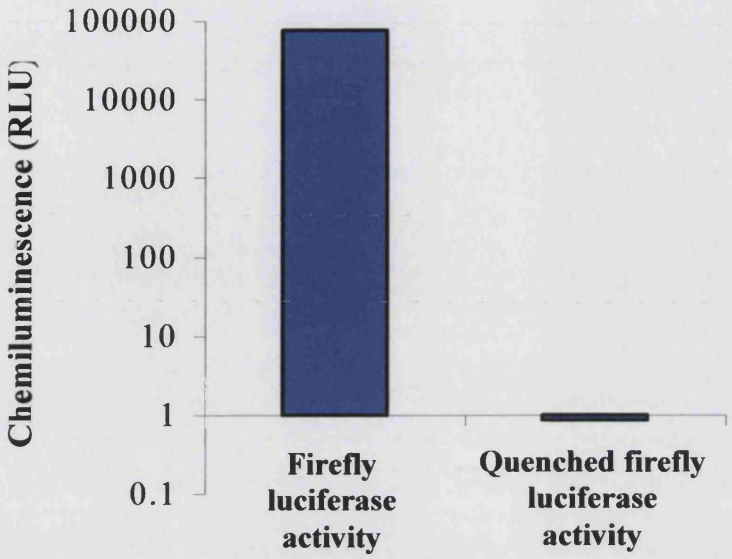


Figure 4.5 Measurement of luciferase activities before and after the addition of 'Stop and Glo'™ reagent.

The graph in Figure 4.5, pg 155, shows that the 'Stop and Glo'™ reagent did successfully quench the firefly luciferase-catalysed reaction. Firefly luciferase luminescence was quenched by greater than 5 orders of magnitude.

4.4.3 Choice of Renilla luciferase internal control vector

Cotransfection of vectors that contain strong promoter/enhancer elements can potentially lead to reductions in reporter gene expression due to the competition for transcription factors between the promoters of the co-transfected vectors (Farr & Roman, 1992). The occurrence and magnitude of the *trans* effects is dependent both on the activity of the genetic regulatory elements present in the co-transfected vectors, and the relative ratio of experimental and control vectors. Since the firefly luciferase expression directed by the HIV-2 LTR within the firefly luciferase reporter vector was to be normalised by the value given by Renilla luciferase expression directed by the internal control reporter vector, it was essential to minimise any potential *trans* effect between the promoters within each cotransfected vector. If these effects were not reduced, *Renilla* luciferase expression directed by the internal control reporter might be down regulated, resulting in an inflated HIV-2 LTR activity value following normalisation. Conversely, if the *trans* effects between the two promoters negatively affected the HIV-2 promoter within

the firefly luciferase reporter vector, the activity value for the experimental HIV-2 LTR would be artificially reduced following normalisation.

Three *Renilla* luciferase reporter vectors were available for use as internal control reporter vectors for cotransfection with the firefly luciferase vector, details of each are summarised below:

pRL-CMV – this vector contains the CMV immediate-early enhancer/promoter region, which provides strong, constitutive expression of *Renilla* luciferase in a variety of cell types (Appendix II, Figure II.4, pg 351).

pRL-TK – this vector contains the herpes simplex virus thymidine kinase (HSV-TK) promoter region upstream of the *Renilla* luciferase gene (Appendix II, Figure II.3, pg 350). The HSV-TK promoter provides low-level, constitutive expression in cells of both embryonal and mature mammalian tissues (Wagner *et al.*, 1985, Stewart *et al.*, 1987).

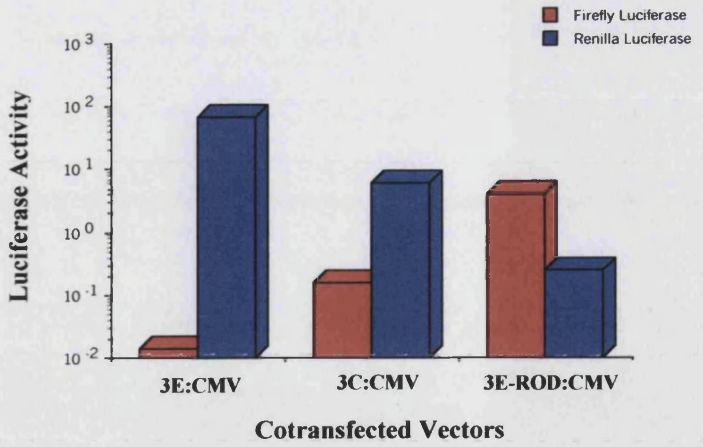
pRL-SV40 – this vector contains the SV40 early enhancer/promoter region, which provides strong, constitutive expression of *Renilla* luciferase in a variety of cell types (Appendix II, Figure II.5, pg 352).

In order to select the luciferase reporter vector pairing that experienced the lowest *trans* effects when co-transfected, the *trans* effects between the promoters in each of the three control reporter vectors and the HIV-2 $pROD_{10}$ LTR that had been inserted into the experimental firefly luciferase vector pGL3E - (3E-ROD) were assessed and compared. Three sets of cotransfections were performed with each internal control vector. Firstly, HEK 293 cells were co-transfected in 12 well dishes with 1.25μg of pRL-TK, pRL-SV40, or pRL-CMV and 1.25μg of pGL3E according to the protocol described in section 2.2.4.4a. Since the firefly luciferase vector pGL3E lacked a promoter region, the level of *Renilla* luciferase gene expression detected in each cotransfection gave a measure of the activity of the corresponding HSV-TK, CMV, or SV40 promoter, in the absence of any *trans* effects.

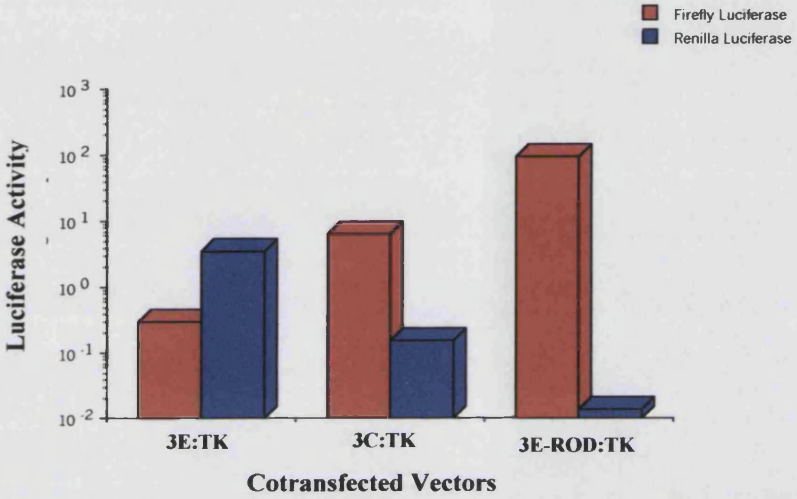
Secondly, HEK 293 cells were cotransfected in 12 well dishes with 1.25μg of pRL-TK, pRL-SV40, or pRL-CMV and 1.25μg of pGL3C. This vector differed from pGL3E in that it contained the early SV40 promoter region as well as the SV40 early enhancer (Appendix II, Figure II.2, pg 349), co-transfecting this vector with each of the internal control vectors provided a positive control for firefly luciferase expression.

Finally, HEK 293 cells were co-transfected in 12 well dishes with 1.25μg of pRL-TK, pRL-SV40, or pRL-CMV and 1.25μg of 3E-ROD (pGL3E containing the HIV-2 $pROD_{10}$ LTR). The level of *Renilla* and firefly luciferase expression detected in this set of cotransfections gave an indication of the activity of the HSV-TK, CMV, SV40 and HIV-2 $pROD_{10}$ promoter regions in the presence of any *trans* effects between the co-transfected promoters. All transfections were performed in duplicate at an internal control vector:experimental reporter vector ratio of 1:1. Following 48h of incubation, cells were harvested and lysed according to the protocol described in section 2.2.4.5 and the firefly and *Renilla* luciferase enzyme activity measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. Results are presented as the mean firefly and *Renilla* luciferase activity for each reporter vector cotransfection, with each mean activity value being calculated from a duplicate cotransfection. The mean firefly and *Renilla* luciferase activities derived in each of the three sets of cotransfections, i.e. internal control vector : pGL3E, internal control vector : pGL3C, and internal control vector : 3E-ROD have been plotted for each of the three internal control vectors (Figure 4.6, pg 158).

(A)
Cotransfections performed with pRL-CMV



(B)
Cotransfections performed with pRL-TK



(C)
Cotransfections performed with pRL-SV40

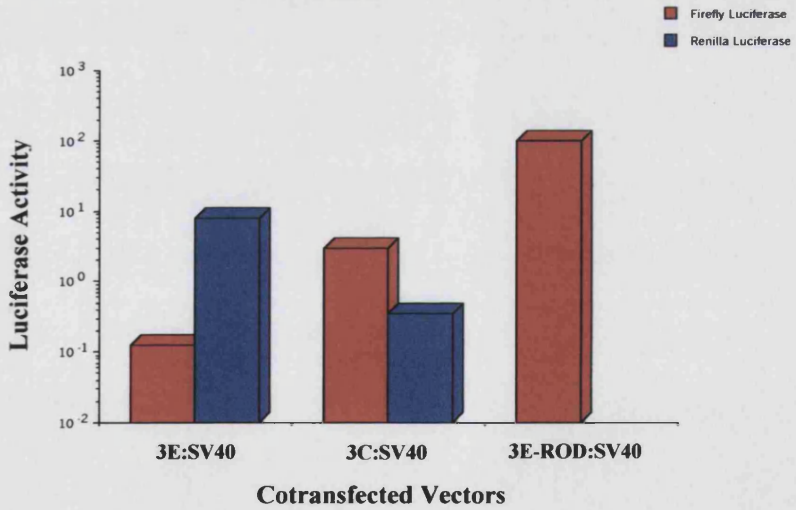


Fig 4.6 Comparison of firefly and *Renilla* luciferase expression in three sets of cotransfections performed with each of three internal control *Renilla* luciferase vectors. (A) Columns 1-2: firefly and *Renilla* luciferase expression from a pGL3E : pRL-CMV cotransfection, 3-4: firefly and *Renilla* luciferase expression from a pGL3C : pRL-CMV cotransfection, 5-6: firefly and *Renilla* luciferase expression from a 3E-ROD : pRL-CMV cotransfection. (B) columns 1-2: firefly and *Renilla* luciferase expression from a pGL3E : pRL-TK cotransfection, 3-4: firefly and *Renilla* luciferase expression from a pGL3C : pRL-TK cotransfection, 5-6: firefly and *Renilla* luciferase expression from a 3E-ROD : pRL-TK cotransfection. (C) columns 1-2: firefly and *Renilla* luciferase expression from a pGL3E : pRL-SV40 cotransfection, 3-4: firefly and *Renilla* luciferase expression from a pGL3C : pRL-SV40 cotransfection, 5-6: firefly and *Renilla* luciferase expression from a 3E-ROD : pRL-SV40 cotransfection. All transfections were performed in duplicate at an internal control vector:experimental reporter vector ratio of 1:1.

4.4.3.1 Trans effects experienced by the promoters contained within the internal control vectors

The blue columns of the graphs shown in Figure 4.6, pg 158, represent the mean *Renilla* luciferase expression corresponding to the activity of the CMV, HSV-TK, and SV40 promoters when co-transfected with the firefly luciferase vectors pGL3E, pGL3C, and 3E-ROD (Figure 4.6, pg 158). In addition, the red columns of the graphs show firefly luciferase expression directed by the enhancer or promoter regions within the pGL3E, pGL3C, and 3E-ROD vectors when cotransfected with each of the three internal control reporter vectors (Figure 4.6, pg 158).

Analyses of the *Renilla* luciferase expression directed by the promoters within the three internal control vectors in each of the three sets of cotransfections revealed an inversely proportional association between the activity of the promoters directing the expression of *Renilla* luciferase and the activity of the promoter regions directing expression of firefly luciferase. As firefly luciferase expression increased, as observed in the cotransfections performed with pGL3C or 3E-ROD when compared to the pGL3E cotransfections (Figure 4.6 A), B) and C), pg 158), *Renilla* luciferase expression directed by the corresponding cotransfected internal control vector was seen to decrease. This association was observed when the cotransfections were performed with pRL-CMV, pRL-TK, and pRL-SV40. In each of the three sets of cotransfections the highest level of *Renilla* luciferase activity was observed when the internal control vector was co-transfected with pGL3E, the reporter vector directing the lowest level of firefly luciferase expression. Taken together these results indicated that the CMV, HSV-TK, and SV40 promoters within the Renilla luciferase internal control vectors pRL-CMV, pRL-TK, and pRL-SV40 respectively, experienced considerable *trans* effects when co-transfected with a firefly luciferase reporter vector containing either the SV40 promoter, in the case of pGL3C or, more importantly, the HIV-2 $pROD_{10}$ LTR within 3E-ROD.

By dividing the *Renilla* luciferase expression value measured for each internal control vector in the internal control:pGL3E cotransfection, by the *Renilla* luciferase expression value measured for each internal control vector in the internal control:3E-ROD cotransfection, the fold decrease in internal control promoter activity observed following cotransfection with the reporter vector containing the HIV-2 $pROD_{10}$ LTR was calculated for the CMV, HSV-TK, and SV40 promoters. In Figure 4.7, pg 160, the graph demonstrates that the activity of the SV40 promoter was most affected following cotransfection with 3E-ROD. The *trans* effects between the SV40 promoter and the HIV-2 $pROD_{10}$ LTR resulted in a 790-fold reduction in SV40 promoter activity. Conversely, the results suggested that the HSV-TK

promoter region within pRL-TK experienced the least *trans* effects when co-transfected with 3E-ROD, with its activity being reduced approximately 258-fold below its activity when co-transfected with promoterless pGL3E.

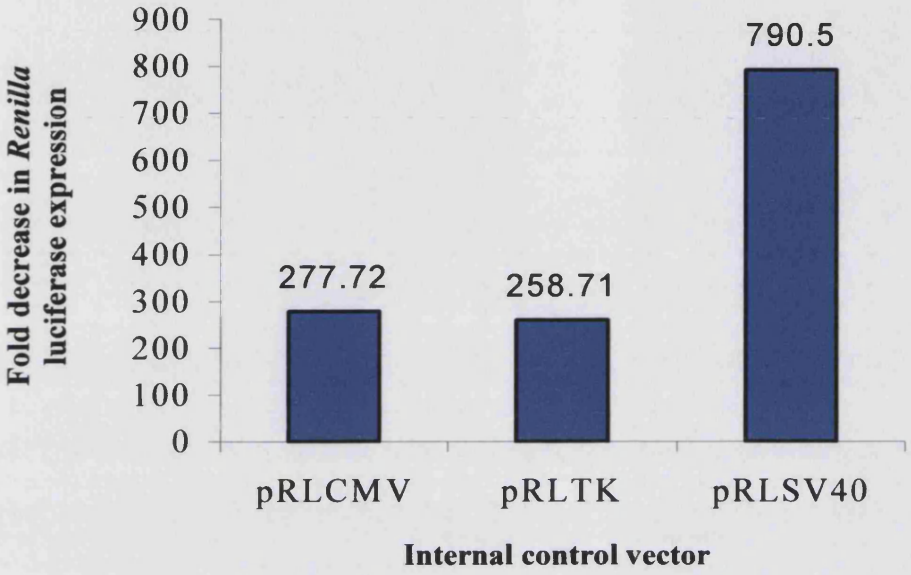


Figure 4.7 Comparison of down regulation experienced by the CMV, HSV-TK, and SV40 promoters within each of three *Renilla* luciferase internal control reporter vectors, when cotransfected with the firefly luciferase reporter vector pGL3E containing an HIV-2 $pROD_{10}$ LTR upstream of the firefly luciferase gene, at a ratio of 1:1.

4.4.3.2 Trans effects experienced by the HIV-2 $pROD_{10}$ LTR within the pGL3E vector

Analyses of the level of firefly luciferase expression measured in each of the three sets of cotransfections (see Figure 4.6 A), B), and C), pg 158) revealed that regardless of which internal control vector it was co-transfected with, the HIV-2 $pROD_{10}$ LTR inserted into the pGL3E reporter vector (3E-ROD) directed significantly higher levels of firefly luciferase expression than the promoterless pGL3E reporter vector, or the SV40 promoter within the pGL3C reporter vector (Figure 4.6, pg 158). The difference between the level of firefly luciferase expression directed by pGL3E and that directed by 3E-ROD is taken a measure of the increase in transcriptional activity directed by the pGL3E vector upon insertion of the HIV-2 $pROD_{10}$ LTR. In the absence of any *trans* effects between the $pROD_{10}$ LTR and the promoter within the co-transfected internal control vector, the increase in activity between the pGL3E and 3E-ROD firefly luciferase values would be the same whether co-transfected with pRL-CMV, pRL-TK, or pRL-SV40. In Figure 4.8, pg 161, the graph showing the calculated increases between the pGL3E and 3E-ROD firefly

luciferase values when co-transfected with pRL-CMV, pRL-TK, and pRL-SV40 showed that this was not the case. The largest difference between the two values occurred when cotransfections were performed using the pRL-SV40 vector. The 795-fold increase in activity of the pGL3E vector possessing an HIV-2 $pROD_{10}$ LTR was significantly higher than the increases of 280.8-fold and 310.5-fold directed by the same vector in the presence of the CMV and HSV-TK promoters respectively. This result demonstrated that the activity of the HIV-2 $pROD_{10}$ LTR was adversely affected least when cotransfected with pRL-SV40, the reporter vector containing the SV40 promoter.

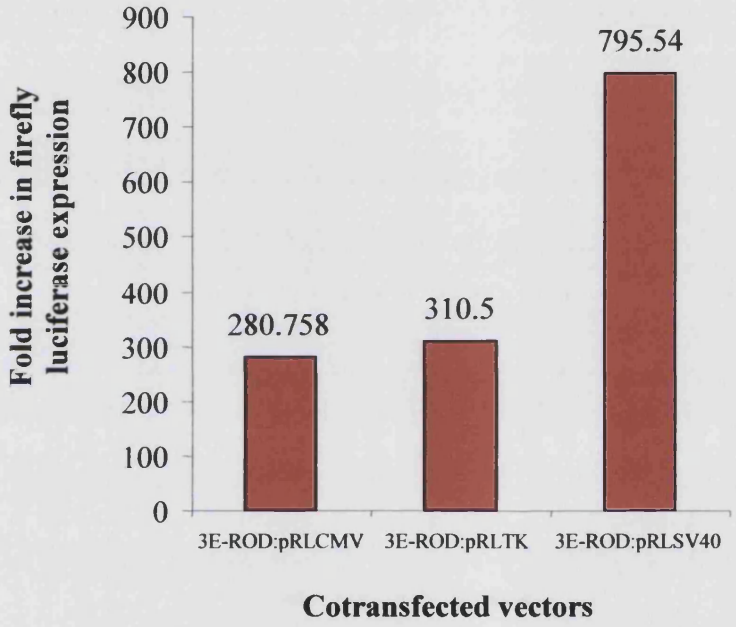


Figure 4.8 Comparison of the fold increase in firefly luciferase expression between the pGL3E vector lacking a promoter and the 3E-ROD vector containing the HIV-2 $pROD_{10}$ LTR, when cotransfected with each of three Renilla luciferase internal control reporter vectors, pRL-CMV, pRL-TK, and pRL-SV40, at a ratio of 1:1.

The promoters within an optimal internal control vector – experimental reporter vector pairing would not experience any *trans* effects when cotransfected. While our results indicated that the HIV-2 $pROD_{10}$ LTR appeared to experience the lowest *trans* effects when cotransfected with pRL-SV40 (Figure 4.8, pg 161), the SV40 promoter within this vector appeared to experience the highest *trans* effects of the internal controls tested when co-transfected with 3E-ROD (Figure 4.7, pg 160). This could potentially lead to problems with *Renilla* luciferase detection, affecting the ability of pRL-SV40 to act as an efficient and reproducible internal control vector. The results shown in Figure 4.6, pg 158, and Figure 4.7, pg 160,

demonstrated that the HSV-TK promoter contained in pRL-TK experienced the least *trans* effects of all the internal control vectors when co-transfected with 3E-ROD. Moreover, second to pRL-SV40, pRL-TK exerted the lowest *trans* effects upon the HIV-2 $pROD_{10}$ LTR (Figure 4.8, pg 161). Based on these findings the pRL-TK vector was selected for use as the internal control vector in all further cotransfection experiments.

4.4.4 Selection of the cotransfection ratio of pRL-TK : pGL3E-HIV-2 LTR

A previous study demonstrating the *trans* effects experienced by the promoters contained within co-transfected vectors indicated that such effects might be reduced by increasing the ratio between the co-transfected vectors (Farr & Roman, 1992). In order to determine the optimal cotransfection ratio of experimental reporter vector DNA to internal control vector DNA, the *trans* effects between the HIV-2 $pROD_{10}$ LTR and the HSV-TK promoter in the firefly and *Renilla* luciferase reporter vectors respectively, were assessed over a range of different cotransfection ratios. HEK293 cells were co-transfected in 12 well dishes with different ratios of 3E-ROD and pRL-TK DNA according to the protocol described in section 2.2.4.4a. Exact 3E-ROD - pRL-TK ratios are given in the legend of Figure 4.9, pg 163. Following 48h, cells were harvested and lysed according to the protocol described in section 2.2.4.5 and the firefly and *Renilla* luciferase enzyme activity measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. Firefly and *Renilla* luciferase values measured at each ratio were plotted and are shown in Figure 4.9, pg 163. The results shown in Figure 4.9, pg 163, indicate an inversely proportional relationship between the amount of pRL-TK within the transfection mixture and the promoter activity of the HIV-2 $pROD_{10}$ LTR. The amount of 3E-ROD within each of the cotransfection mixtures was kept constant, while the level of internal control *Renilla* luciferase reporter vector, pRL-TK was reduced by dilution. If the *trans* effects experienced by the two promoters were the same at each ratio, the level of firefly luciferase expression directed by the HIV-2 $pROD_{10}$ LTR would be constant. The results shown in Figure 4.9, pg 163, demonstrated that this was not the case. As the level of pRL-TK within the cotransfection mixture was reduced, with increasing pRL-TK:3E-ROD ratios, the level of firefly luciferase gene expression directed by the HIV-2 $pROD_{10}$ LTR increased. The highest level of firefly luciferase expression directed by the HIV-2 $pROD_{10}$ LTR was observed at the highest dilution of pRL-

TK, 1:50. These data suggested that negative *trans* effects between the two promoters were being experienced and could be reduced by increasing the ratio between the two reporter vectors.

Although optimal firefly luciferase gene expression was obtained at the pRL-TK:3E-ROD ratio of 1:50, levels of *Renilla* luciferase expression at this ratio, and ratios of 1:40 and 1:20, were at the lower end of detectability by the dual-luciferase assay. The optimal ratio of cotransfection has to be a compromise between a reduction of *trans* effects between the promoters, and the quantity of vector required to give a reproducible and detectable level of reporter gene expression. Therefore all subsequent cotransfections were performed at a pRL-TK:3E-ROD ratio of 1:10. This ratio provided some reduction in the *trans* effects experienced between the two promoters, but still ensured detectable levels of reporter gene expression.

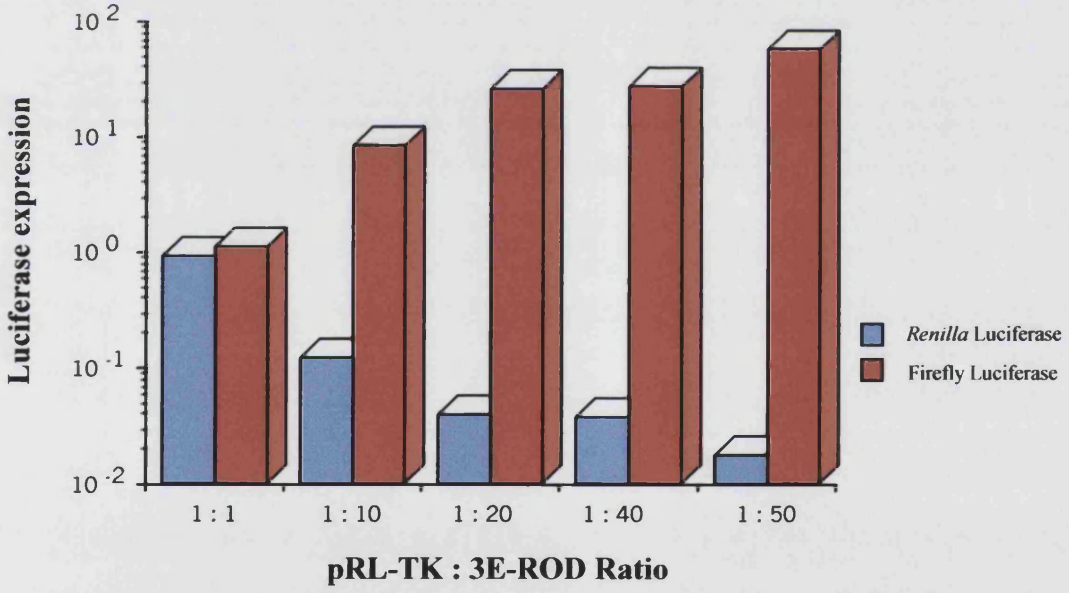


Figure 4.9 Comparison of firefly and *Renilla* luciferase expression at different cotransfection ratios. Columns 1-2 represent firefly and *Renilla* luciferase expression directed when 1.25µg 3E-ROD and 1.25µg of pRL-TK (2.5µg total DNA, 1:1 pRL-TK: 3E-ROD ratio) were cotransfected into HEK293 cells.
Columns 3-4 represent firefly and *Renilla* luciferase expression directed when 1.25µg 3E-ROD, 0.125µg pRL-TK, and 1.125µg of HSDNA (2.5µg total DNA, 1:10 pRL-TK: 3E-ROD ratio) were cotransfected.
Columns 5-6 represent firefly and *Renilla* luciferase expression directed when 1.25µg 3E-ROD, 12.5ng pRL-TK, and 1.2375µg of HSDNA (2.5µg total DNA, 1:20 pRL-TK: 3E-ROD ratio) were cotransfected.
Columns 7-8 represent firefly and *Renilla* luciferase expression directed when 1.25µg 3E-ROD, 1.25ng pRL-TK, and 1.249µg of HSDNA (2.5µg total DNA, 1:40 pRL-TK: 3E-ROD ratio) were cotransfected.
Columns 9-10 represent firefly and *Renilla* luciferase expression directed when 1.25µg 3E-ROD, 0.125ng pRL-TK, and 1.249µg of HSDNA (2.5µg total DNA, 1:50 pRL-TK: 3E-ROD ratio) were cotransfected.

4.4.5 Repeatability of the dual-luciferase detection assay

In order to determine the repeatability of the dual-luciferase detection assay, the lysates from ten different LTR clone cotransfections (performed and prepared according to the protocols described in sections 2.2.4.4a and 2.2.4.5) were assayed in duplicate using the dual-luciferase assay (see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. The normalised firefly luciferase values (normalised = firefly luciferase value divided by the *Renilla* luciferase value measured in the same well) for each of the ten lysates were plotted against their duplicate values in order to assess the variation between wells assayed in duplicate. The high correlation coefficient obtained ($R^2 = 0.9888$) when regression analysis was performed on this data set indicated that very little variation was observed between the two normalised luciferase values produced from the duplicate detection assays (Figure 4.10, pg 165). Therefore, samples were assayed in singlicate only from this point on.

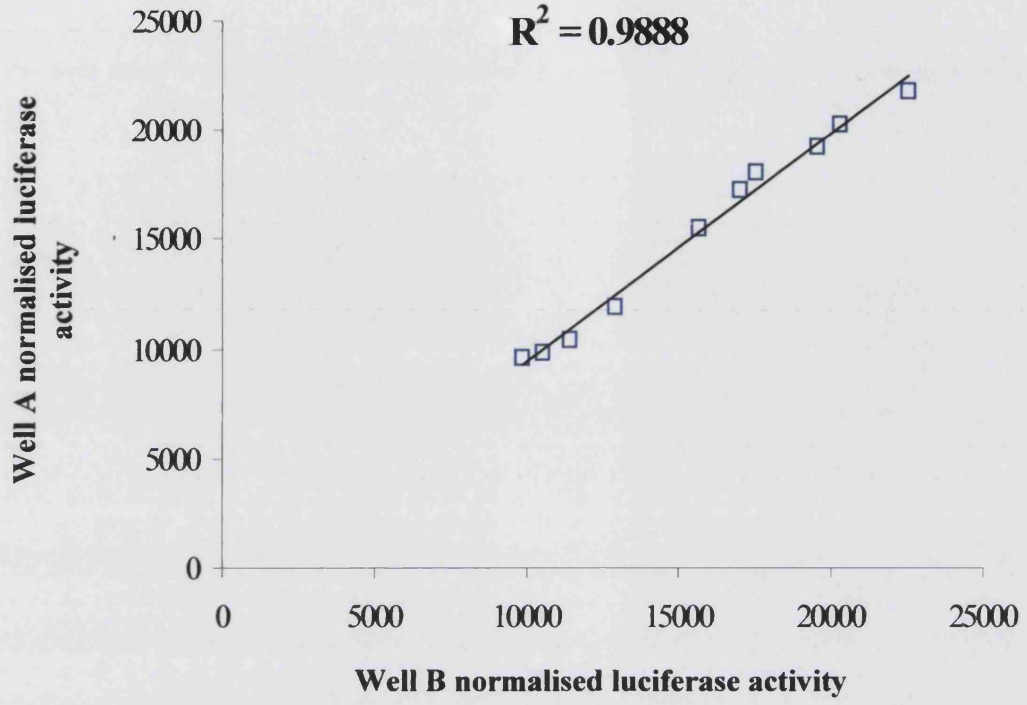


Figure 4.10 Comparison of normalised firefly luciferase chemiluminescence from the same transfected cell lysate measured in duplicate wells (n=10). Regression analysis of well A chemiluminescent values plotted against well B chemiluminescent values is shown.

4.5 Transient cotransfection of Jurkat and THP-1 cells

Transient transfection allows the rapid analysis of regulatory elements through reporter gene expression. When cells are transiently transfected, the DNA is introduced into the nucleus of the cell but does not integrate into the chromosomes. This means that many copies of the gene of interest are present, leading to high levels of expressed protein. Every cell type has a characteristic set of requirements for optimal introduction of vector DNA. Several different transient transfection techniques are currently available; the technique used is an important factor in determining the transfection efficiency of the cell type. In order to obtain the most favourable cotransfection conditions for the Jurkat and THP-1 cell lines in which the comparative HIV-2 LTR studies were to be performed, a number of optimisation experiments were carried out.

4.5.1 Transient cotransfection of nonadherent cell lines

Transient transfection protocols utilising DEAE-dextran- or calcium phosphate-mediated gene transfer have been widely used to transfect non-adherent cell lines such as the Jurkat T-cell-like cell-line, and the THP-1 monocyte-like cell-line. Both techniques produce a chemical environment that induces the attachment of DNA to the cell surface; the DNA is then endocytosed and transported to the cell nucleus. Although these techniques are relatively simple and reproducible, both induce significant cytotoxic effects and require transfected cells to be temporarily grown in reduced serum conditions. This often leads to slow recovery rates for transfected cells, resulting in reduced levels of protein expression. Non-adherent cultures are also commonly transfected by electroporation, a technique that uses an electric field to open up pores in the cell, through which the DNA diffuses. A significant disadvantage of this technique is that approximately 5-fold greater quantities of DNA and cells are needed than in either of the above two methods.

At the time of this study two new transfection reagents became available that claimed to result in outstanding transfection efficiencies while being significantly less toxic to cells than the previously described transfection reagents (Qiagen). The first was Superfect (Qiagen), this reagent is a specifically designed activated dendrimer possessing branches, radiating in a spherical structure that terminate in charged amino groups. When complexed with DNA the Superfect complex possesses a net positive charge which allow it to bind to negatively charged receptors (e.g. sialylated glycoproteins) on the surface of eukaryotic cells. Endocytosis and transport to the cellular nucleus follows. The second reagent, called

Effectene (Qiagen), is a non-liposomal lipid reagent that transfers DNA into mammalian cells in combination with a special DNA-condensing reagent, named 'enhancer'. Since this study called for large numbers of transfection experiments to be performed with the non-adherent Jurkat and THP-1 cell-lines, we opted to use the new reagents, based upon their enhanced transfection efficiencies and the reduced potential for cytotoxic effects in comparison to the more classical transfection techniques.

Optimal transfection is influenced by several parameters, including cell density, cell type, the amount of DNA transfected, and the transfection reagent to DNA ratio. Jurkat and THP-1 cell-lines were transfected using both Superfect and Effectene reagents in an attempt to establish which reagent directed the highest transfection efficiency in which cell-line.

4.5.2 Optimisation of transient cotransfections using Superfect and Effectene transfection reagents.

In order to determine which transfection reagent directed the greatest transfection efficiency in which cell-line four sets of cotransfections were performed. Each cell-line was cotransfected using each transfection reagent, over a range of DNA concentrations at a number of different DNA:transfection reagent ratios, such that the conditions, in addition to the transfection reagent, giving the greatest transfection efficiency in each cell-line could be determined simultaneously. The level of reporter gene expression directed in the optimal Superfect transfection for Jurkat and for THP-1 cells was compared to the level of reporter gene expression directed in the optimal Effectene transfection in each cell-line. In this way, it was possible to determine whether Superfect or Effectene directed the greatest transfection efficiencies in each cell-line, and at what DNA concentration and DNA:transfection reagent ratio.

4.5.2.1 Transfection of Jurkat cells using Superfect

When following the Superfect protocol, Jurkat cells were cotransfected in 6-well dishes, each well was seeded with approximately 2 x 10^6 cells. Each of three different amounts of DNA (1μg, 2μg, and 4μg of 3E-ROD and pRL-TK DNA at a ratio of 10:1) were transfected into Jurkat cells using three different DNA:Superfect reagent ratios (1:2, 1:5 and 1:10); according to the protocol described in section 2.2.4.6a. Exact amounts of DNA and Superfect are given in the legend of Figure 4.11, pg 169. Cotransfections were performed in duplicate. Following 48h of incubation, cells were harvested and lysed according the protocol described in section 2.2.4.7, and the firefly and *Renilla* luciferase enzyme activity measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount)

(see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. Each firefly luciferase activity measurement corresponding to the enzyme expression directed by the HIV-2 $pROD_{10}$ LTR was normalised for transfection variation by dividing it by the corresponding cotransfected *Renilla* luciferase activity measurement. Results are presented as the mean normalised firefly luciferase activity for each experiment, with each value being calculated from a duplicate cotransfection (Figure 4.11, pg 169).

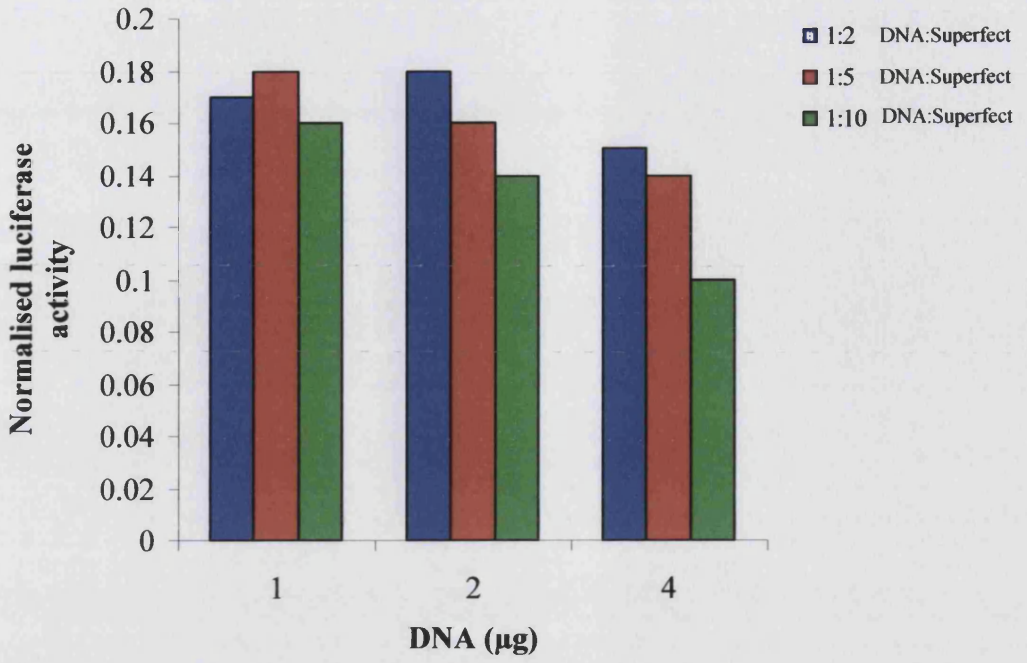


Figure 4.11 Evaluation of the optimal Superfect (Qiagen) cotransfection conditions in the Jurkat cell line. Columns 1-3 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 900ng 3E-ROD, 90ng of pRL-TK, and 10ng HSDNA (1µg total DNA) were transfected with increasing levels of transfection reagent, 1.8µl (1:2), 3.6µl (1:5) and 9µl (1:10) of Superfect respectively.
Columns 4-6 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 1800ng 3E-ROD, 180ng pRL-TK, and 20ng HSDNA (2µg total DNA) were transfected with increasing levels of transfection reagent, 3.6µl (1:2), 9µl (1:5) and 18µl 1:10) of Superfect respectively.
Columns 7-9 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 3600ng 3E-ROD, 360ng pRL-TK, and 40ng HSDNA (4µg total DNA) were transfected with increasing levels of transfection reagent, 7.2µl (1:2), 18µl (1:5) and 36µl (1:10) of Superfect respectively. Cotransfections were performed in duplicate.

The results in Figure 4.11, pg 169, indicated the existence of two inversely proportional relationships. The first was between DNA:Superfect ratio and reporter gene expression, the second between DNA concentration and reporter gene expression. Regardless of whether cotransfections were performed with a total DNA concentration of 1, 2, or 4µg, levels of reporter gene expression were seen to decrease as the ratio between the DNA and Superfect reagent in the cotransfection mixture increased. At each of the three DNA concentrations (1, 2, and 4µg), cotransfections performed at the DNA:Superfect ratio of 1:10 directed the lowest levels of normalised luciferase activity while, (with the exception of the cotransfection performed with 1µg DNA, DNA:Superfect ratio 1:2) cotransfections performed at ratio of 1:2 directed the highest levels of normalised luciferase activities (Figure 4.11, columns 3, 6, and 9, vs. columns 1, 4, and 7, pg 169).

In addition, at each ratio a gradual decrease in normalised firefly luciferase activity was noted as the DNA concentration within the cotransfection mixture was increased from 1μg to 2μg to 4μg (Figure 4.11, columns 2 vs. 5 vs. 8, and columns 3 vs. 6 vs. 9, pg 169). Each of the three cotransfections performed with 1μg of DNA (Figure 4.11, columns 1-3, pg 169) directed higher levels of normalised luciferase activity than the three cotransfections performed with 4μg of DNA (Figure 4.11, columns 7-9, pg 169). The highest levels of normalised luciferase activity were observed when either 1μg of DNA was transfected with Superfect at a 1:5 dilution, or 2μg of DNA was transfected with Superfect at a 1:2 dilution. The lowest levels of activity were found in cotransfections performed with the highest concentration of DNA (4μg) and DNA:Superfect ratio (1:10).

4.5.2.2 Transfection of Jurkat cells using Effectene

When following the Effectene protocol, Jurkat cells were cotransfected in 6-well dishes, each well was seeded with approximately 2 x 10^6 cells. As with the Superfect optimisation, Jurkats were transfected with three different amounts of DNA (0.2μg, 0.4μg, and 0.8μg of 3E-ROD and pRL-TK DNA at a ratio of 10:1) using three different DNA:Effectene reagent ratios (1:10, 1:25 and 1:50); according to the protocol described in section 2.2.4.6b. The 'enhancer' reagent volume within each cotransfection mixture was adjusted with respect to the Effectene reagent. Exact amounts of DNA, Effectene, and 'enhancer' are given in the legend of Figure 4.12, pg 171. Cotransfections were performed in duplicate. Following 48h of incubation, cells were harvested and lysed according the protocol described in section 2.2.4.7, and the firefly and *Renilla* luciferase enzyme activity measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. Each firefly luciferase activity measurement was normalised for transfection variation by dividing it by the corresponding cotransfected *Renilla* luciferase activity measurement. Results are presented as the mean normalised firefly luciferase activity for each experiment, with each value being calculated from a duplicate cotransfection (Figure 4.12, pg 171).

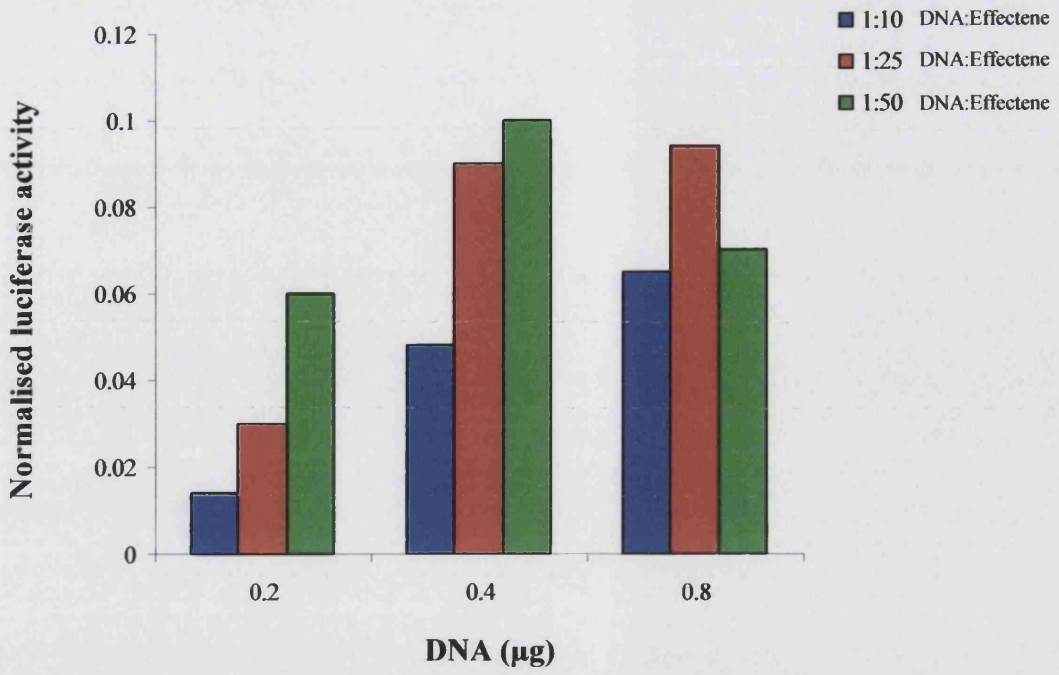


Figure 4.12 Evaluation of the optimal Effectene (Qiagen) cotransfection conditions in the Jurkat cell line. Columns 1-3 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 180ng 3E-ROD, 18ng pRL-TK, 2ng HSDNA (0.2µg total DNA) were transfected with 1.6µl of 'enhancer' and increasing levels of transfection reagent, 2µl (1:10), 5µl (1:25) and 10µl (1:50) of Effectene respectively. Columns 4-6 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 360ng 3E-ROD, 36ng pRL-TK, and 4ng HSDNA (0.4µg total DNA) were transfected with 3.2µl of 'enhancer' and increasing levels of transfection reagent, 4µl (1:10), 10µl (1:25) and 20µl 1:50) of Effectene respectively. Columns 7-9 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 720ng 3E-ROD, 72ng pRL-TK, and 8ng HSDNA (0.8µg total DNA) were transfected with 6.4µl of 'enhancer' and increasing levels of transfection reagent, 8µl (1:10), 20µl (1:25) and 40µl (1:50) of Effectene respectively. Cotransfections were performed in duplicate.

The results shown in Figure 4.12, pg 171, indicated the existence of two directly proportional relationships. In direct contrast to the Superfect data (Figure 4.11, pg 169) normalised firefly luciferase activity was found to increase as the DNA:Effectene ratio was increased within the cotransfection mixture. At each of the three DNA concentrations, levels of normalised luciferase activities directed in cotransfections performed at a DNA:Effectene ratio of 1:50 were higher than activities directed in cotransfections performed at the 1:10 ratio (with the exception of the 0.8µg 1:50 cotransfection, column 9).

In addition, levels of normalised firefly luciferase activity were found to increase at each of the DNA:Effectene ratios as the total amount of cotransfected DNA in the cotransfection mixture was increased. At each ratio normalised firefly luciferase activities were seen to increase as DNA concentration was increased from 0.2µg to 0.4µg to 0.8µg (Figure 4.12, columns 1 vs. 4 vs. 7, and

columns 2 vs. 5 vs. 8, pg 171). Following this it was expected that the cotransfection performed with the greatest amount of DNA (0.8µg) at the highest DNA:Effectene ratio (1:50) would direct the highest level of normalised firefly luciferase activity, however the activity measured for this cotransfection was unexpectedly low. Cotransfections performed with either 0.8µg of DNA at a DNA:Effectene ratio of 1:25, or with 0.4µg of DNA at a DNA:Effectene ratio of 1:50 were found to direct the highest levels of normalised firefly luciferase activity.

When the normalised firefly luciferase activity value obtained under optimal conditions for the Superfect transfection reagent (Figure 4.11, column 4, pg 169) was compared to the normalised firefly luciferase activity obtained in the optimal Effectene reagent transfection in Jurkat cells (Figure 4.12, column 6, pg 171) it became clear that Superfect directed the highest transfection efficiency (Figure 4.13, pg 172). The cotransfection performed using 2µg of DNA at a 1:2 ratio with Superfect gave a normalised firefly luciferase activity value that was approximately 2-fold higher than the activity directed when reporter vectors were cotransfected with 0.4µg of DNA at a DNA:Effectene ratio of 1:50 (Figure 4.13, pg 172). Therefore, all further transfection experiments using Jurkat cells were performed in 6-well dishes, using 2µg of DNA at a 1:2 ratio with Superfect.

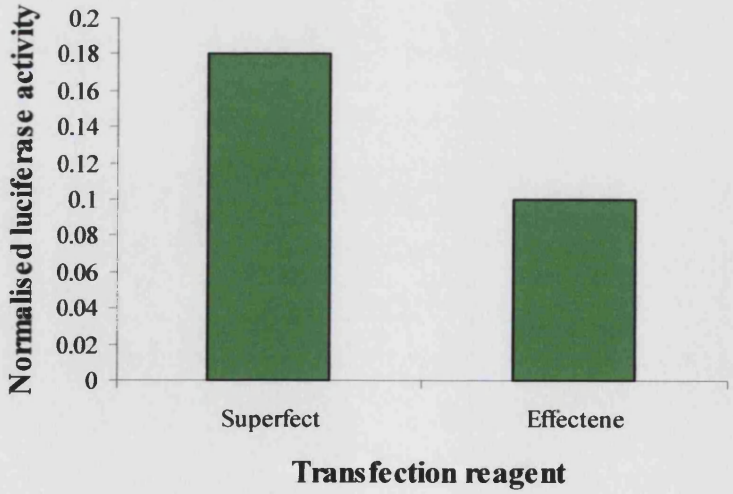


Figure 4.13 Comparison of normalised firefly luciferase expression directed by 3E-ROD in the Jurkat cell-line when cotransfected in optimal conditions using Superfect or Effectene as the transfection reagent.

4.5.2.3 Transfection of THP-1 cells using Superfect

2 x 10^6 THP-1 cells were cotransfected in 6-well dishes using a similar pipetting scheme to the one described for Jurkat cells (see section 4.5.2.1), in order to determine the optimum transfection conditions for THP-1 cells when using the Superfect reagent. Cells were cotransfected with three different amounts of DNA (1μg, 2μg, and 4μg of 3E-ROD and pRL-TK DNA at a ratio of 10:1) at three different DNA:Superfect reagent ratios (1:2, 1:5 and 1:10); according to the protocol described in section 2.2.4.6a. Exact amounts of DNA and Superfect are given in the legend of Figure 4.14, pg 174. Cotransfections were performed in duplicate. Following 48h of incubation, cells were harvested and lysed according the protocol described in section 2.2.4.7, and the firefly and *Renilla* luciferase enzyme activity measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. Firefly luciferase activity was normalised for transfection variation by dividing it by the corresponding cotransfected *Renilla* luciferase activity measurement. Results are presented as the mean normalised firefly luciferase activity for each experiment, with each value being calculated from a duplicate cotransfection (Figure 4.14, pg 174).

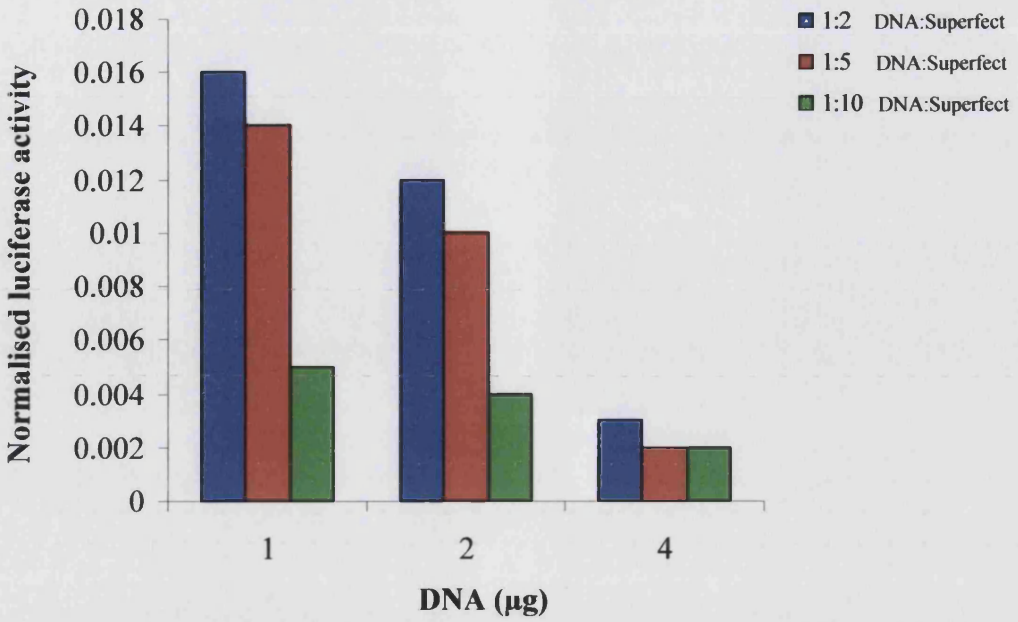


Figure 4.14 Evaluation of the optimal Superfect (Qiagen) cotransfection conditions in the THP-1 cell line. Columns 1-3 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 900ng 3E-ROD, 90ng of pRL-TK, and 10ng HSDNA (1µg total DNA) were transfected with increasing levels of transfection reagent, 1.8µl (1:2), 3.6µl (1:5) and 9µl (1:10) of Superfect respectively.
Columns 4-6 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 1800ng 3E-ROD, 180ng pRL-TK, and 20ng HSDNA (2µg total DNA) were transfected with increasing levels of transfection reagent, 3.6µl (1:2), 9µl (1:5) and 18µl 1:10) of Superfect respectively.
Columns 7-9 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 3600ng 3E-ROD, 360ng pRL-TK, and 40ng HSDNA (4µg total DNA) were transfected with increasing levels of transfection reagent, 7.2µl (1:2), 18µl (1:5) and 36µl (1:10) of Superfect respectively. Cotransfections were performed in duplicate.

The results in Figure 4.14, pg 174, demonstrated that THP-1 cells behaved similarly to Jurkat cells when cotransfected with the Superfect reagent, although the effects of changing the DNA concentration and DNA:Superfect ratio on transfection efficiency appeared to be more significant. The two inversely proportional relationships that were observed in Figure 4.11, pg 169, were also observed in Figure 4.14, pg 174, but appeared more pronounced. Normalised firefly luciferase activities were found to decrease significantly at each of the three DNA concentrations when the DNA:Superfect ratio was increased (Figure 4.14, columns 1, 4, and 7 vs. columns 3, 6, and 9, pg 174). Cotransfections performed at a DNA:Superfect ratio of 1:10 directed significantly lower levels of normalised luciferase activity than cotransfections performed at a ratio of 1:2 or 1:5, irrespective of the total cotransfected DNA concentration (i.e. 1, 2 or 4µg). At each of the three DNA concentrations the DNA:Superfect ratio of 1:2 directed the highest levels of normalised luciferase activity.

In addition, the level of normalised firefly luciferase activity at each DNA:Superfect ratio was found to decrease significantly as the DNA concentration within the cotransfection mixture was increased from 1μg to 2μg to 4μg (Figure 4.11, columns 2 vs. 5 vs. 8, and columns 3 vs. 6 vs. 9, pg 169). Normalised firefly luciferase activities from the cotransfections performed with 4μg of DNA (Figure 4.14, columns 7-9, pg 174) directed significantly lower activities than the activities from the three cotransfections performed with 1μg of DNA (Figure 4.14, columns 1-3, pg 174). The highest levels of normalised luciferase activity were observed when 1μg of DNA was transfected with Superfect at a 1:2 dilution. As observed with Jurkat cells the lowest levels of activity were found in cotransfections performed with the highest concentration of DNA (4μg) and DNA:Superfect ratio (1:10). However, even at optimal cotransfection conditions overall levels of normalised firefly luciferase activity directed by THP-1 cells were much lower.

4.5.2.4 Transfection of THP-1 cells using Effectene

2 x 10^6 THP-1 cells were cotransfected in 6-well dishes using a similar pipetting scheme to the one described for Jurkat cells (see section 4.5.2.2), in order to determine the optimum transfection conditions for THP-1 cells when using the Effectene reagent. Cells were cotransfected with three different amounts of DNA (0.2μg, 0.4μg and 0.8μg) using three different DNA:Effectene reagent ratios (1:10, 1:25 and 1:50); according to the protocol described in section 2.2.4.6b. The 'enhancer' reagent volume within each cotransfection mixture was adjusted with respect to the Effectene reagent. Exact amounts of DNA, Effectene, and 'enhancer' are given in the legend of Figure 4.15, pg 177. Cotransfections were performed in duplicate. Following 48h of incubation, cells were harvested and lysed according the protocol described in section 2.2.4.7, and the firefly and *Renilla* luciferase enzyme activity measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activity was assessed (section 2.2.5.2) and subtracted from all corresponding measurements. Each firefly luciferase activity measurement was normalised for transfection variation by dividing it by the corresponding cotransfected *Renilla* luciferase activity measurement. Results are presented as the mean normalised firefly luciferase activity for each duplicate cotransfection (Figure 4.15, pg 177).

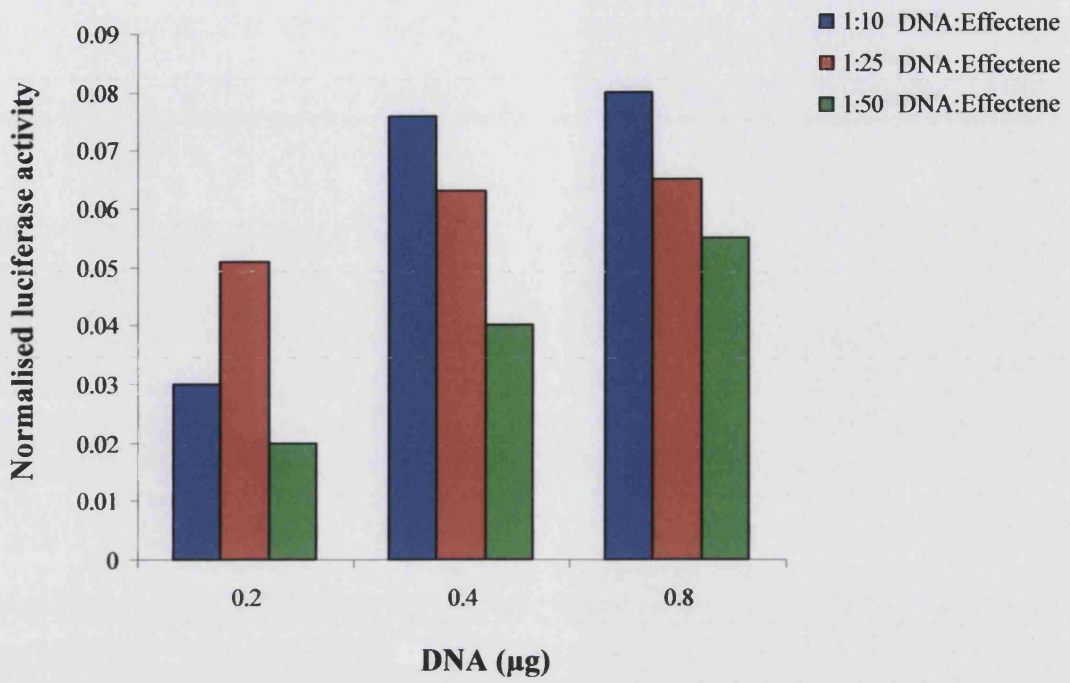


Figure 4.15 Evaluation of the optimum reporter DNA:Effectene transfection ratio in the THP-1 cell line. Columns 1-3 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 180ng 3E-ROD, 18ng pRL-TK, 2ng HSDNA (0.2µg total DNA) were transfected with 1.6µl of 'enhancer' and increasing levels of transfection reagent, 2µl (1:10), 5µl (1:25) and 10µl (1:50) of Effectene respectively. Columns 4-6 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 360ng 3E-ROD, 36ng pRL-TK, and 4ng HSDNA (0.4µg total DNA) were transfected with 3.2µl of 'enhancer' and increasing levels of transfection reagent, 4µl (1:10), 10µl (1:25) and 20µl 1:50) of Effectene respectively. Columns 7-9 represent normalised firefly luciferase activity directed by the $pROD_{10}LTR$ when 720ng 3E-ROD, 72ng pRL-TK, and 8ng HSDNA (0.8µg total DNA) were transfected with 6.4µl of 'enhancer' and increasing levels of transfection reagent, 8µl (1:10), 20µl (1:25) and 40µl (1:50) of Effectene respectively. Cotransfections were performed in duplicate.

The results shown in Figure 4.15, pg 177, indicated the existence of an inversely proportional relationship between normalised firefly luciferase activity and DNA:Effectene ratio, and a directly proportional relationship between normalised firefly luciferase activity and the total amount of DNA transfected.

At each of the three DNA concentrations, levels of normalised luciferase activity were seen to decrease as the ratio between DNA and Effectene increased. Cotransfections performed at a DNA:Effectene ratio of 1:50 directed lower levels of normalised firefly luciferase activities than those performed with the same amount of DNA but at a DNA:Effectene ratio of 1:10. This finding was in contrast to the data generated from the Effectene transfection of Jurkat cells, where, at each of the three DNA concentrations, cotransfections performed at a DNA:Effectene ratio of 1:50 directed the highest normalised firefly luciferase activities (Figure 4.12, pg 171). However, a similar pattern was observed when THP-1 cells were cotransfected with the Superfect reagent (Figure 4.14, pg 174).

Similar to the data generated when Jurkat cells were cotransfected with Effectene (Figure 4.12, pg 171), levels of normalised firefly luciferase activity were found to increase at each of the DNA:Effectene ratios as the total amount of DNA within the cotransfection mixture was increased from 0.2μg to 0.4μg to 0.8μg (Figure 4.15, columns 1 vs. 4 vs. 7, and columns 2 vs. 5 vs. 8, pg 177). Following this, cotransfections performed with the greatest amount of DNA (0.8μg) at the lowest DNA:Effectene ratio, 1:10, were found to direct the highest levels of normalised firefly luciferase activity.

When the normalised firefly luciferase activity value obtained under optimal conditions for the Effectene transfection reagent (Figure 4.15, column 7, pg 177) was compared to the normalised firefly luciferase activity obtained in the optimal Superfect reagent transfection in THP-1 cells (Figure 4.14, column 1, pg 174) it became clear that Effectene directed the highest transfection efficiency (Figure 4.16, pg 178). The cotransfection performed using 0.8μg of DNA at a 1:10 ratio with Effectene gave a normalised firefly luciferase activity value that was 5-fold higher than the activity directed when reporter vectors were cotransfected with 1μg of DNA at a DNA:Superfect ratio of 1:2 (Figure 4.16, pg 178). Therefore, all further transfection experiments using THP-1 cells were performed in 6-well dishes, using 0.8μg of DNA at a 1:10 ratio with Effectene.

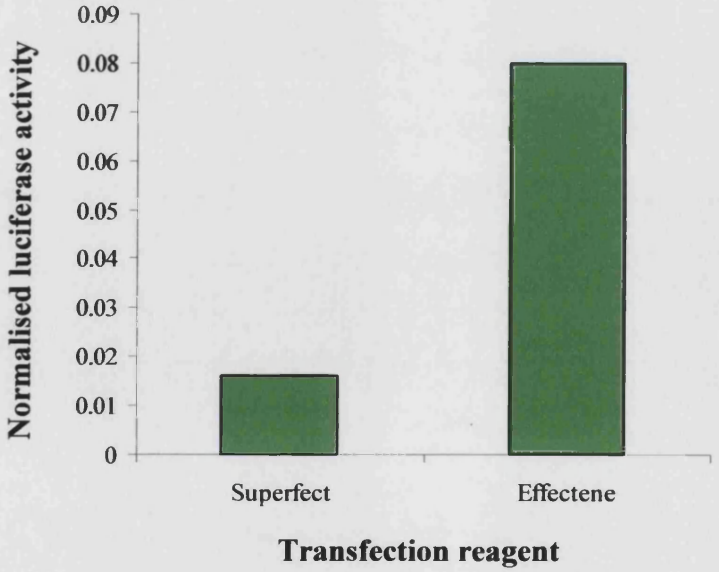


Figure 4.16 Comparison of normalised firefly luciferase expression directed by 3E-ROD in the THP-1 cell-line when cotransfected in optimal conditions using Superfect or Effectene as the transfection reagent.

4.5.3 Further optimisation of transient cotransfections using Superfect and Effectene transfection reagents.

4.5.3.1 Optimal cell density for greatest transfection efficiency

In view of the fact that large numbers of transfection experiments were to be performed during the course of this study, minimisation of the quantities of DNA, cells, and reagents used in each experiment was paramount. The optimisation of the concentrations of DNA and transfection reagents have been described above, however, the manufacturer's instructions accompanying the Superfect and Effectene reagents indicated that each reagent could work over a range of cell densities. In an attempt to reduce the number of cell required to perform each experiment cotransfections were performed in six well dishes seeded with 1 x 10^6 cells/well rather than 2 x 10^6 cells/well using the optimised conditions described for the Superfect transfection of Jurkat cells and the Effectene transfection of THP-1 cells. Cotransfections were performed in duplicate and a mean normalised firefly luciferase activity for each cotransfection was determined (see sections 2.2.4.6, and 2.2.5). Results are presented as the mean normalised firefly luciferase activity for each duplicate cotransfection performed (Figure 4.17, pg 179).

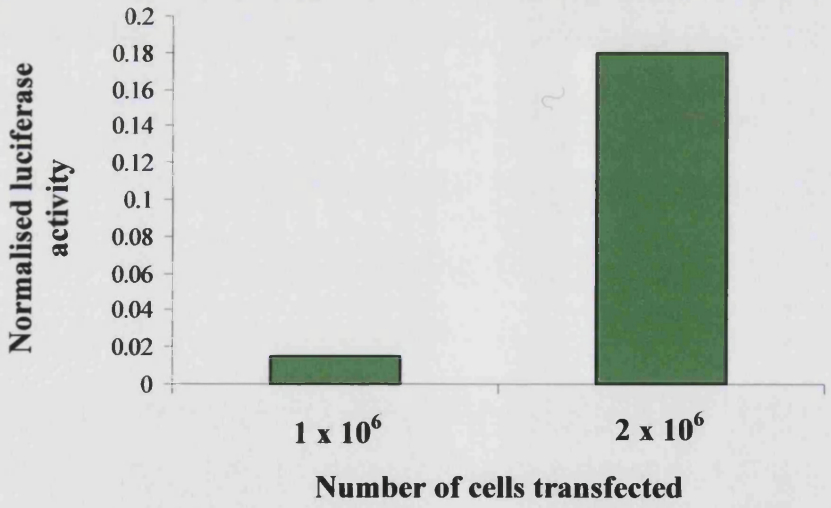


Figure 4.17 Comparison of normalised firefly luciferase activity detected when 1 x 10^6 and 2 x 10^6 Jurkat cells were cotransfected with 1800ng 3E-ROD and 200ng of pRL-TK (2µg total DNA) using 3.6µl of Superfect (DNA:Superfect ratio, 1:2) according to the protocol described in section 2.2.4.6.a.

At a cell density of 1 x 10^6 cells/well, levels of normalised firefly luciferase were reduced by approximately 12-fold in the Jurkat cell transfection (Figure 4.17, pg 179), and to undetectable levels within the THP-1 cell data set (data not shown). Based on these results both cell lines were transfected at a cell density of 2 x 10^6 cells/well.

4.5.3.2 Optimal post-cotransfection conditions for greatest transfection efficiency

The greatest transfection efficiency in Jurkat cells had been achieved when the Superfect transfection reagent was used to cotransfect cells with 2µg of DNA at a DNA:Superfect ratio of 1:2. According to the manufacturers (Qiagen), transfection efficiency could be further enhanced under optimal Superfect transfection conditions by adding a washing step to the protocol 2 hours post-transfection. 2 x 10^6 Jurkat cells were cotransfected according to the conditions described in section 2.2.4.6a. Two hours post-transfection cells were pelleted and resuspended in fresh complete medium (Appendix I). 48 hours post-transfection cells were harvested and lysed (2.2.4.7) and normalised firefly luciferase activities were determined (2.2.5). Cotransfections were performed in duplicate. Results are presented as the mean normalised firefly luciferase activity for each duplicate cotransfection (Figure 4.18, pg 180).

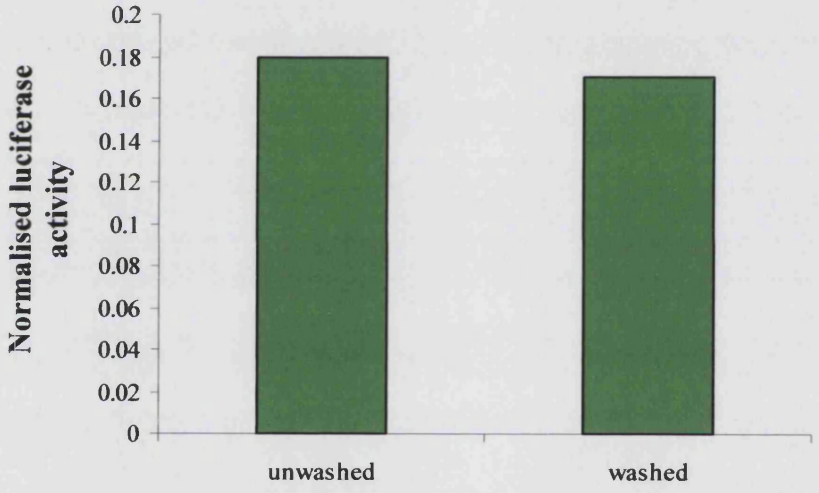


Figure 4.18 Comparison of normalised firefly luciferase activity detected following a cotransfection procedure including a washing step 2 hours post-transfection, and a cotransfection procedure without a washing step 2 hours post-transfection. Cotransfections were performed using 2 x 10^6 Jurkat cells under optimised conditions (section 4.5.2.1).

The results obtained following this procedure indicated that the additional washing step did not greatly improve the final level of gene expression (Figure 4.18, pg 180). Slightly lower levels of reporter expression were noted in the transfected cells that had been washed. Consequently, the extra washing step was not added to the final Superfect transfection protocol for Jurkat cells.

4.6 Reproducibility of cotransfection

4.6.1 Variation of luciferase expression within a duplicate cotransfection

In order to assess the variation in firefly and *Renilla* luciferase expression from duplicate cotransfections the normalised luciferase activities from the cotransfections of ten different HIV-2 LTR clones were plotted against the values obtained in the duplicate cotransfections. The high correlation coefficient obtained (R^2 = 0.9848) when regression analysis was performed on this data set indicated that very little variation in firefly or *Renilla* luciferase expression occurred between wells transfected in duplicate (Figure 4.19, pg 181). Therefore, the normalised luciferase activity values obtained from duplicate cotransfections could be pooled without severely altering the final normalised luciferase expression value calculated to represent each HIV-2 LTR clone.

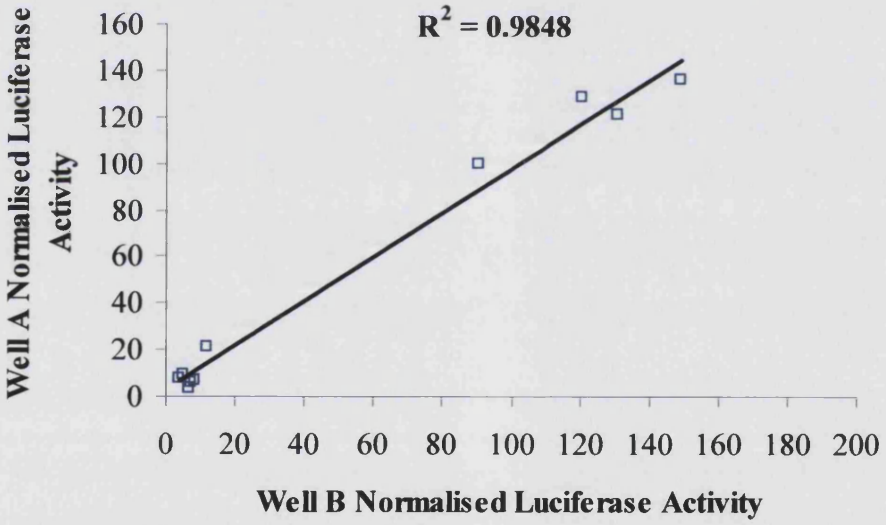


Figure 4.19 Comparison of normalised firefly luciferase activity from duplicate cotransfections. This graph compares the normalised firefly luciferase activity expressed from each of the two replicates of a duplicate cotransfection, in ten instances (n=10). Ten different HIV-2 LTR clones were cotransfected into Jurkat cells in duplicate with the internal control vector pRL-TK at a ratio of 10:1. Regression analysis of well A chemiluminescent values plotted against well B chemiluminescent values are shown.

4.6.2 Variation of luciferase expression in cotransfections performed at two different time points

Plotting the normalised luciferase activities from the cotransfection of ten different HIV-2 LTR clones against the values taken from the same cotransfections performed one week later assessed the reproducibility of cotransfection (Figure 4.20, pg 182). The average normalised luciferase activity values taken from a triplicate cotransfection of each of the ten different HIV-2 LTR clones were plotted. The high correlation coefficient obtained (R^2 = 0.8796) when regression analysis was performed on this data set indicated that very little variation was observed in the level of firefly and *Renilla* luciferase expression directed by the two reporter vectors when co-transfected at two different time points. This enabled the comparison of results obtained from LTR clones co-transfected at different time points, and ensured that any differences in activity resulting from cotransfection experiments could be noted and would not be as a result of the variation in the cotransfection efficiency of different experiments.

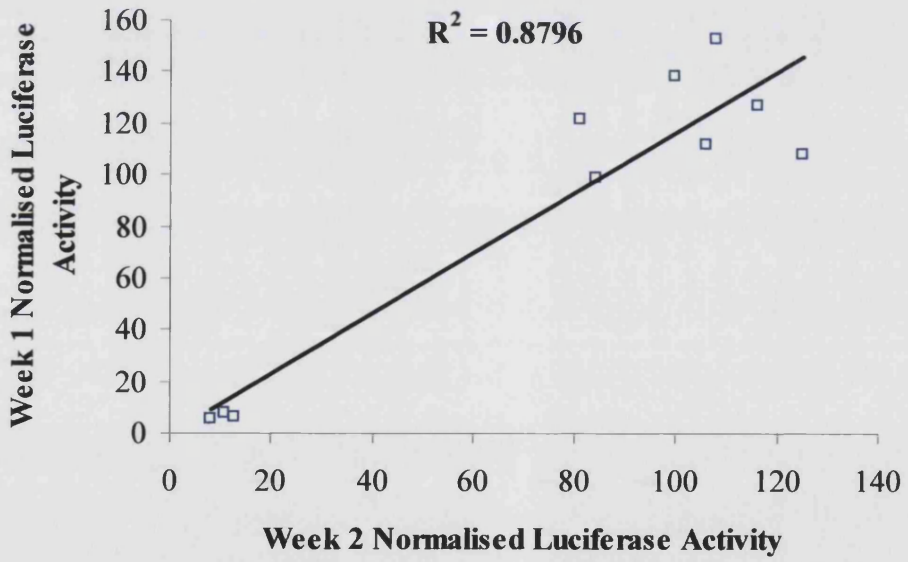


Figure 4.20 Evaluation of cotransfection reproducibility. This graph compares normalised firefly luciferase activity expressed from the same clone cotransfected into Jurkat cells with the internal control vector pRL-TK, at two different time points (n=10). Regression analysis of well A chemiluminescent values plotted against well B chemiluminescent values is shown.

4.7 Summary and discussion

This chapter focused on the optimisation of an efficient HIV-2 LTR cloning strategy, the development of a sensitive and reproducible reporter detection assay, and the evaluation of optimal cotransfection conditions within the Jurkat and THP-1 cell-lines, for the assessment of the activities of HIV-2 LTRs taken from a wide variety of viral sources.

The dual-luciferase reporter assay (Promega Ltd) offered several advantages over other available reporter assays for the analyses of genetic regulatory elements, the most important being the inclusion of an internal control vector, the assays' extremely broad linear range of detection, and the speed at which the detection assay could be performed. The inclusion of an internal control vector enabled normalisation of the expression directed by the experimental promoter within the firefly luciferase reporter vector, effectively eliminating variability introduced into the assay by differences in transfection efficiencies and the number and health of the transfected cells. The broad linear range of the detection assay enabled the analysis of both large and small changes in promoter activity, while the speed of the detection assay ensured that large-scale comparative studies of HIV-2 LTR activity could be performed rapidly.

Selection of the experimental firefly luciferase reporter vector pGL3E (Promega Ltd), a reporter vector containing an SV40 enhancer but lacking a promoter region, ensured high levels of reporter gene expression in mammalian cell lines. Optimisation of linearisation and cloning techniques facilitated the rapid and efficient cloning of PCR amplified HIV-2 LTR sequences from diverse viral DNA sources into pGL3E, and resulted in the routine achievement of cloning efficiencies of up to 80%. A minimum of five HIV-2 LTR clones were obtained from each member of the CBL series of HIV-2 isolates and each of the six members of the Gambian HIV-2 cohort for use in the comparative analysis of HIV-2 LTR activity (Chapters 5, 6, and 7). In addition, three clones of the HIV-2 $pROD_{10}$ LTR were obtained for use as positive controls in the optimisation experiments outlined in this chapter.

Assessment of the sensitivity of the dual-luciferase detection assay showed a dynamic range spanning at least seven orders of magnitude, from 0.1pg to 10μg (50pg/ml and 5mg/ml). Analysis of the stability of the signal emitted by the reaction catalysed by the firefly luciferase enzyme revealed only a negligible drop in activity over the entire dynamic range. Since the luciferase detection assay is performed in 96-well plates, which the luminometer reads in less than six minutes, any drop in activity that occurs in the wells over the course of plate reading will be negligible. Examination of the repeatability of the detection assay by regression analysis revealed very little variation in the normalised luciferase activity between

wells assayed in duplicate. The sensitivity and repeatability of the luciferase detection assay ensured that reporter gene expression directed by promoters of varying activities could be measured in singlicate, accurately and reproducibly.

Evaluation of the optimal internal control vector – experimental reporter vector cotransfection pairing revealed that the promoters within each of the three internal control vectors and the HIV-2 $pROD_{10}$ LTR inserted into pGL3E, experienced significant *trans* effects when cotransfected. Similar findings have been described by Farr and coworkers (Farr & Roman, 1992) when investigating the activity of the upstream regulatory region (URR) of a human papillomavirus (HPV6-W50). When cotransfected with a CAT reporter vector containing the URR and the SV40 early promoter the β-gal activity directed by the β-gal internal control vector pRSV β-gal was consistently found to be 7 – 10-fold lower than upon cotransfection with a CAT reporter vector lacking the URR (Farr & Roman, 1992). Of the three internal control vectors, the activity of the HSV-TK promoter contained within pRL-TK was least affected when cotransfected with the pGL3E reporter vector containing the HIV-2 $pROD_{10}$ LTR, while the activity of the SV40 promoter was down-regulated the most. Compared to the CMV and SV40 promoters the TK promoter region is a weaker promoter region, directing low level rather than strong constitutive expression of *Renilla* expression in a variety of cell types. Therefore, as a promoter it might compete less for transcription factors with the HIV-2 LTR and explain the lower *trans* effects experienced by this promoter when cotransfected with 3E-ROD. Interestingly, the HIV-2 $pROD_{10}$ LTR experienced the lowest *trans* effects when cotransfected with the SV40 and not the TK promoter. An explanation for this finding might lie in the fact that the SV40 promoter was the most downregulated of the three internal control vectors and was the least able therefore, to exert a competitive influence upon the transcription factors within the cell nucleus. However, it is apparent that the relationships between the cotransfected promoters and the competition for the transcription factors that bind to them are very complex. Since the TK promoter experienced the lowest *trans* effects when cotransfected with 3E-ROD clone, and the HIV-2 $pROD_{10}$ LTR experienced the second lowest *trans* effects when cotransfected with the vector containing the TK promoter, pRL-TK was selected for use as the internal control vector in all subsequent experiments.

Optimisation experiments performed to determine the optimal cotransfection ratio between the internal control vector and the experimental reporter vector revealed that the *trans* effects experienced between 3E-ROD and pRL-TK could be reduced. The highest level of firefly luciferase expression was measured from the cotransfection performed at a ratio of 3E-ROD to pRL-TK of 50:1. Detection of *Renilla*

luciferase expression at this ratio however, was only just above background. The optimal ratio of cotransfection has to be a compromise between a reduction of *trans* effects between the promoters, and the quantity of internal control vector required to give a reproducible and detectable level of reporter gene expression. Therefore, all subsequent cotransfections were performed at a 3E-ROD:pRL-TK: ratio of 10:1. This ratio provided a 10-fold reduction in the *trans* effects between the two promoters, and ensured detectable levels of reporter gene expression.

Once optimised the dual-luciferase assay system was used to determine the optimal cotransfection conditions for both the Jurkat and THP-1 cell lines. Variation in DNA concentration and the ratio between transfection reagent and DNA had significant effects upon the transfection efficiency of both cell lines. When transfecting both Jurkat and THP-1 cells levels of normalised firefly luciferase activity were found to decrease as the level of Superfect reagent and total amount of DNA within the cotransfection mixture was increased. The reduction in reporter gene expression was a result of an increase in cell growth inhibition and cell death as the concentration of the two components was increased in the mixture, and could be verified visually as well as by dual luciferase assay. The reverse was true in Jurkat cells transfected with Effectene where levels of normalised firefly luciferase activity were found to increase as the concentration of Effectene and total DNA within the cotransfection mixture was increased, indicating that of the two reagents Superfect was more toxic than Effectene at higher concentrations in Jurkat and THP-1 cells. Despite this, under optimal conditions cotransfection of Jurkat cells using a low Superfect and total DNA concentration yielded 2-fold higher levels of normalised firefly luciferase activity than those measured from an optimal Effectene cotransfection of Jurkat cells. Superfect was chosen therefore as the reagent to transfect Jurkat cells with in all subsequent experiments. In THP-1 cells, the toxicity of the Superfect-DNA complex was so great that even at low concentrations the normalised firefly luciferase activities measured from these cotransfections were lower than most of the cotransfections performed using Effectene. The sensitivity of THP-1 cells to the toxicity of transfection reagents was also underlined by the fact that, at each of the three DNA concentrations, cotransfections performed with the 1:50 DNA:Effectene ratio directed lower normalised firefly luciferase activities than the activities measured in the cotransfections performed at a DNA:Effectene ratio of 1:10. Under optimal cotransfection conditions the Effectene transfection technique yielded 5-fold higher levels of normalised firefly luciferase activity in THP-1 cells than the optimal Superfect technique for this cell-line, and was subsequently chosen as the transfection reagent for THP-1 cells. Interestingly, in line with the findings

presented in this thesis, Qiagen now indicate in their technical resources manual that the Superfect reagent is not a suitable for the transfection of the THP-1 cell-line.

In an attempt to further optimise the cotransfection procedures for both Jurkat and THP-1 cells the effects of altering the cell density of cotransfection and including an additional washing process post-transfection were investigated. Reporter gene expression was reduced significantly in both Jurkat and THP-1 cells when the number of cells transfected was reduced from 2 x 10^6 cells to 1 x 10^6 cells. It is likely that the reduction in reporter gene expression is due to a smaller number of cells directing reporter gene expression rather than a decrease in the transfection efficiency/DNA uptake of the cells that are present. Following this, all subsequent cotransfections were performed using 2 x 10^6 Jurkat or THP-1 cells. Ultimately optimal conditions for transfection of both cell lines involved a compromise between reporter gene expression rates and cell death. Washing transfected Jurkat cells 2 hours post-transfection did not significantly increase normalised luciferase values or decrease growth inhibition and cell death, suggesting that the process of transfection rather than the presence of the transfection reagents in the cell media had greater effects upon cellular growth. Therefore, the additional washing step was not included in the final Jurkat cotransfection protocol.

Analyses of the reproducibility of cotransfection revealed no significant differences between the levels of normalised firefly luciferase activity taken from each well of a duplicate cotransfection. Furthermore, very little variation was detected between the normalised firefly luciferase values obtained from identical cotransfections performed one week apart. The reproducibility of the cotransfection protocols in addition to the inclusion of an internal control vector ensured that normalised firefly luciferase activity values obtained from replicate cotransfections could be pooled without severely altering the value calculated to represent each HIV-2 LTR clone, and that normalised firefly luciferase activity values obtained from cotransfections carried out at different time points could be reliably and confidently compared.

To conclude, a dual-luciferase reporter assay system capable of rapidly and reproducibly measuring the activity of cloned HIV-2 LTRs taken from diverse viral sources over a dynamic range extending at least seven orders of magnitude has been optimised. The cloning, cotransfection, and detection protocols utilised by the assay system have been developed to enable reproducible analyses of both large and small changes in HIV-2 LTR activity. In addition, optimisation of the system has ensured that results from assays performed at different time points can be reliably compared. Both cloning, transfection, and reporter assay protocols developed in this thesis are also being used to assess HIV-2 LTR activity in

samples derived from the Caio cohort, Guinea Bissau (Okarofor *et al.,* unpublished), and HIV-1 LTR activity in brain and lymph-node samples (Malik *et al.,* unpublished).

Chapter 5.

Functional Characterisation of the HIV-2 LTR in the CBL Series of HIV-2 Isolates.

5.1 Introduction

Following the identification of a second retrovirus in healthy female commercial sex-workers in Senegal in 1985 (Barin *et al.*, 1985), and the subsequent isolation of HIV-2 a year later, from patients with AIDS in Guinea Bissau and Cape Verde Islands (Clavel *et al.*, 1986a, Brun-Vezinet *et al.*, 1987), The Gambia has represented one of the major countries in the region with a significant prevalence of HIV-2 infection. While the prevalence of HIV-2 was very low in The Gambia in 1987 (A.Wilkins, unpublished observations), subsequent seroepidemiological studies have indicated a prevalence of around 1.8% to be present in the general population by the beginning of the 1990s (Wilkins *et al.*, 1991a, Wilkins *et al.*, 1991b). HIV-2 is therefore thought to have been a relatively recent introduction into The Gambia, probably from nearby or neighbouring countries such as Senegal and Guinea Bissau (Schim van der Loeff & Aaby, 1999). Detailed studies concerning the natural history, epidemiology, and significance of HIV-2 in The Gambia have been conducted at the MRC laboratories, Fajara, located on the Atlantic coast of the country (Wilkins *et al.*, 1991a, Wilkins *et al.*, 1991b, Whittle *et al.*, 1994, Ariyoshi *et al.*, 1996, Whittle *et al.*, 1998a).

Clinical, immunological and virological studies have led to the characterisation of HIV-2 infections, indicating that while HIV-2 is capable of acting as a pathogenic retrovirus, the vast majority of HIV-2 infections are associated with a less aggressive disease course (Poulsen *et al.*, 1989, Pepin *et al.*, 1991b, Ricard *et al.*, 1994, Norrgren *et al.*, 1995, Ariyoshi *et al.*, 1996, Lisse *et al.*, 1996). In general, patients infected with HIV-2 experience a much more prolonged asymptomatic period of infection than had been previously characterised in HIV-1 infection (Pepin *et al.*, 1991b, Whittle *et al.*, 1994, Jaffar *et al.*, 1997). Comparative epidemiological evidence has confirmed that progression to immunosuppresion and AIDS in the majority of cases of HIV-2 infection is slower than for HIV-1 (Adjorlolo-Johnson *et al.*, 1994, Kanki *et al.*, 1994, Marlink *et al.*, 1994, Whittle *et al.*, 1994, Poulsen *et al.*, 1997). At the MRC hospital in The Gambia, individuals with more progressive HIV-2 infections have been identified, along with those who are asymptomatic (Poulsen *et al.*, 1989, Ricard *et al.*, 1994, Norrgren *et al.*, 1995, Poulsen *et al.*, 1997, Ariyoshi *et al.*, 1998). Thus a broad spectrum of HIV-2 infections has been identified. The first isolates

to be made from this country were made in collaboration with the Chester Beatty Laboratories and were designated the CBL series of isolates (Schultz *et al.*, 1990).

A number of recent studies have chartered the association between levels of viral replication or viremia circulating within the peripheral blood of HIV-2 infected individuals, and the rate of progression to disease (Simon *et al.*, 1993, Berry *et al.*, 1998, Ariyoshi *et al.*, 2000, Popper *et al.*, 2000). A study by Ariyoshi *et al* has shown that base-line HIV-2 RNA load significantly predicts the rate of disease progression as determined by CD4+ decline or by death (Ariyoshi *et al.*, 2000). HIV-2 infected individuals with high plasma viral loads were found to lose $CD4^+$ cells rapidly and died quickly. In contrast, HIV-2 infected individuals with low plasma viral loads showed better or no disease progression. Several groups including Ariyoshi and colleagues have suggested that viral turnover is elevated within HIV-2 infected rapid progressors but restricted in HIV-2 infected individuals who show little or no disease progression (Berry *et al.*, 1998, Grassly *et al.*, 1998, Ariyoshi *et al.*, 2000, Popper *et al.*, 2000).

The factors involved in determining low-level viral replication in the majority of HIV-2 individuals, particularly during the asymptomatic period, and high viral replication in others, have yet to be determined

To assess the relationship between LTR function, replication capacity and, disease progression status in HIV-2 infection, a series of studies were undertaken. Initially, these were performed with the CBL series of isolates CBL-20 – 24, derived from Gambian HIV-2-infected individuals (Shultz *et al.*, 1990). Previous characterisation of the biological and molecular variability of these isolates revealed different replicative phenotypes *in vitro* and suggested that, as with HIV-1, a relationship existed between the *in vitro* cytopathogenicity of the different isolates and the clinical disease status of the patient from whom they were originally isolated (Shultz *et al.*, 1990). Studies were then extended to include characterisation of naturally occurring HIV-2 LTR genotypes from two contrasting groups of HIV-2 infected Gambian individuals displaying different rates of disease progression and disease association. The following chapter describes the characterisation of LTR function within the CBL series of HIV-2 isolates. The activity of LTRs cloned from CBL 20 – 24 (described in Chapters 3 and 4) has been measured at basal and Tat-induced levels within two biologically relevant cell-lines, the Jurkat T-cell-like cell line and the monocyte-like THP-1 cell-line, using the optimised dual luciferase reporter assay system (described in

Chapter 4). By comparing the HIV-2 CBL isolates key differences in HIV-2 LTR activity may be revealed.

5.2 Summary of isolate details

Blood samples were taken from 20 anti-HIV-2 seropositive individuals attending the MRC hospital or the clinic for sexually transmitted diseases in Fajara, The Gambia and mononuclear cells were cultured for ten days (all procedures were performed by colleagues in The Gambia). In total, seven HIV-2 strains derived from different disease categories were isolated by colleagues at the Chester Beatty Laboratories, London (Shultz *et al.*, 1990, CBL-20 – CBL-26). We were able to obtain DNA from five of the original seven HIV-2 isolates (CBL-20 - 24) which differ in their growth rates, cytopathogenicity *in vitro*, and sensitivity to neutralising antibodies in patient sera. In addition, there is a close correlation between the *in vitro* cytopathogenicity of the different isolates and the clinical disease status of the individuals from whom they were obtained (Table 5.1, pg 192). The cytopathogenicity and replication of each of the isolates was assessed in several cell lines including peripheral blood mononuclear cells, MOLT4, H9, C8166 and CEM cell lines (by colleagues at the Chester Beatty Laboratories, London) (Schultz *et al.*, 1990). CBL-20 and CBL-21, two highly cytopathic strains, grew rapidly in primary culture and produced high levels of reverse transcriptase within 14 days, and were derived from patients who had progressed to AIDS (Table 5.1, pg 192). In contrast, CBL-22 and CBL-23 took somewhat longer to produce significant reverse transcriptase activity, and took slightly longer to appear in primary culture. CBL-22 was derived from an AIDS-related complex (ARC) patient who went on to die. CBL-23 was derived from an asymptomatic patient whose condition progressed to ARC (Table 5.1, pg 192). CBL-24 was the least cytopathic isolate of the five isolates analysed. This isolate grew only in the MOLT4 cell-line after several attempts, and was derived from an asymptomatic patient.

Isolate	Sex	Age (yr)	Clinical Status at time of isolation	Disease Progression
CBL-20	M	28	AIDS, diarrhoea, wasting, oral thrush	Dead
CBL-21	F	34	AIDS, diarrhoea, wasting	AIDS
CBL-22	M	55	ARC, diarrhoea, wasting, oral thrush	Dead
CBL-23	F	31	Asymptomatic	ARC
CBL-24	F	22	Asymptomatic	PGL

Table 5.1 Clinical characteristics of individuals from whom the HIV-2 CBL isolates were obtained. Multiple LTRs were cloned from each of five members of the CBL series and were used to assess LTR function at the basal and Tat-induced level within the dual luciferase reporter assay system. (Reproduced from Shultz *et al*, 1990).

5.3 Transcriptional activity of HIV-2 isolate-derived LTRs

The role of HIV-2 LTR function in determining replicative capacity and disease progression status was initially examined using LTRs cloned from five members of the CBL series of HIV-2 isolates.

5.3.1 Evaluation of the functional variability of LTRs taken from a single HIV-2 isolate

The variation in basal activity directed by multiple LTRs cloned from a single CBL isolate was initially examined. This experiment provided data on the range of LTR activities present within a viral quasispecies and enabled a judgement to be made about the number of LTR clones required to give a representative view of the LTR activity of each isolate within comparative transfection experiments. CBL-22 was the randomly selected member of the CBL series chosen for this analysis.

Ten CBL-22 LTR clones (produced by the nested LTR PCR and restriction digest cloning procedures described in chapters 2 and 3) were cotransfected into Jurkat cells with *Renilla* luciferase internal control reporter vector pRL-TK at a ratio of 10:1 according to the protocol set out in section 2.2.4.6a. Exact DNA concentrations are given in the legend of Figure 5.1, pg 194. Each LTR clone was cotransfected in triplicate. Following 48h of incubation, cells were harvested, lysed, and basal CBL-22-LTR-directed firefly luciferase expression and RSV-TK promoter-directed *Renilla* luciferase expression was detected from each lysate using the dual-luciferase assay (see section 2.2.5). Background activity was assessed according to section 2.2.5.2 and subtracted from all corresponding measurements.

Three normalised basal activity values were calculated for each CBL-22 LTR clone by dividing the level of firefly luciferase expression by the level of *Renilla* luciferase expression measured in each well of the triplicate cotransfection. This experiment was repeated one week later using fresh Jurkat cells, and a second data set consisting of three normalised luciferase activity values for each of the CBL-22-LTR clones obtained. When tested by one-way ANOVA no significant difference was seen between the first and second data sets obtained one week apart for four CBL-22 LTR clones selected at random ($p=0.15 – 0.53$). Therefore, a mean normalised basal activity value was calculated for each CBL-22 LTR clone by pooling the data from all six replicates (i.e. three normalised basal activity values from each of two triplicate transfections performed at different time points), without significantly altering the final normalised basal activity value representing the activity of the each LTR. Where analyses were subsequently carried out at more than one time, data from all time points were combined. The mean

normalised basal activity values directed by each of the ten LTR clones from the CBL-22 isolate are shown in Figure 5.1, pg 194.

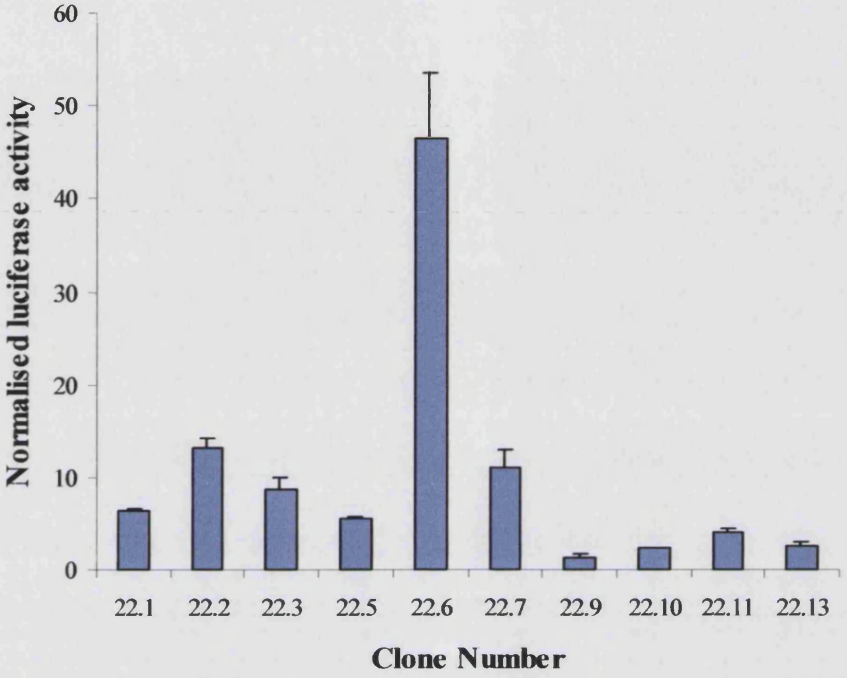


Figure 5.1 Differential basal activity of ten CBL-22 LTR clones (n=1). The cloned LTRs were tested for basal activity in 2 x 10^6 Jurkat cells. Co-transfections were performed with 1.25µg CBL-22 LTR-pGL3E, 0.125µg of pRL-TK, and 0.625µg of HSDNA. Luciferase activities directed by the HIV-2 LTRs were normalised with respect to the *Renilla* luciferase activity using the dual-luciferase assay system (Promega Ltd). Each value represents the mean normalised basal LTR activity of six cotransfections, with each error bar representing the standard deviation of the six individual values.

Each of the ten CBL-22 LTRs cloned into the firefly luciferase plasmid pGL3E were functional at the basal level, the lowest basal activity was directed by clone 22.9, (normalised luciferase activity – 1.36), while the highest basal activity was directed by the LTR clone 22.6 (normalised luciferase activity – 46.53). This represented a 30 – 40-fold range in activity.

When tested by one-way ANOVA, statistically significant differences ($p<0.0001$) were found between the mean normalised basal activities of the ten CBL22 LTR clones. By comparing intra-clonal sampling variation and the variation between the mean basal LTR activities of the LTR clones in the analysis this test demonstrated that the mean basal activity of at least one of the clones was significantly different to the mean basal LTR activities of some or all of the remaining LTR clones in the analysis. This analysis showed that the inter-clone difference in LTR activity was greater than the variation between the six

normalised activity values obtained for each LTR clone from two triplicate transfection experiments, and was unlikely, therefore, to be explained by sampling variation.

Multiple comparisons by Tukey-Kramer tests revealed that the mean normalised basal LTR activity of clone 22.6 was significantly higher ($p<0.0001$) than the mean normalised basal activities of the other nine CBL-22 LTR clones. However, significant differences were also found between the mean normalised basal LTR activities of clone 22.2 and 22.9 ($p=0.0477$), and clone 22.2 and 22.10 ($p=0.0225$), indicating that the significant differences suggested by the initial one way ANOVA above, were not solely generated by the activity of clone 22.6.

Taken together the results indicated that multiple LTR clones would give a more representative view of promoter activity within the viral quasispecies of an isolate than a single LTR clone. Since clones from each of the five CBL isolates were to be transfected in triplicate into both Jurkat and THP-1 cell lines, and were to be assessed for basal and Tat-induced activity within both cell lines, a practicable figure of five clones was chosen to represent the promoter region of each isolate.

5.3.2 Functional analysis of LTRs cloned from five HIV-2 isolates within the Jurkat cell line

HIV-2 LTR functional studies were undertaken with the CBL series of HIV-2 isolates initially in the T-cell like Jurkat cell line. CBL-20 – CBL24 were each represented by five LTR clones, produced as described in chapters 3 and 4. LTR function at both the basal and Tat-induced levels for each of the clones was assessed and compared using the optimised dual luciferase assay (see section 2.2.5).

5.3.2.1 Basal activity of HIV-2 LTRs cloned from the CBL series of HIV-2 isolates within the Jurkat cell-line

The basal activities of the LTRs cloned from HIV-2 isolates CBL-20 to CBL-24 were assessed in Jurkat cells using the Superfect transfection reagent (Qiagen) protocol described in section 2.2.4.6a. Each LTR clone was transfected in triplicate; precise DNA concentrations are given in the legend of Figure 5.2A, pg 198. Luciferase enzyme activity was measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activities were assessed as described in section 2.2.5.2, and subtracted from all corresponding measurements. These cotransfections were performed in parallel to those described in section 5.3.2.2. Results are presented as the mean normalised basal activity for each LTR clone, calculated from the three normalised basal activity values of the triplicate transfection (Figure 5.2A, pg 198), and as a box plot representing the range of the five clonal basal LTR activity values measured for each CBL isolate (Figure 5.2B, pg 198). The median basal LTR activity for each isolate is marked by a line, the box representing the distance between the first and third quartiles, and the leading and trailing 'whiskers' corresponding to the minimum and maximum respectively (Figure 5.2B, pg 198). Differences between LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant.

The results from the Jurkat data set revealed that each of the twenty-five cloned LTRs from the five members of the CBL series of HIV-2 isolates (CBL-20 – CBL24) was able to direct a detectable level of basal activity within the Jurkat cell-line (Figure 5.2A, pg 198), although the LTRs within clones 21.5 and 22.4, taken from CBL 21 and 22 respectively, displayed particularly low basal activities. Significant variation was observed between the basal activities of LTRs cloned from the same CBL isolate and LTRs cloned from different CBL isolates, with basal activities overall ranging from 0.88 to 19.30. In this cell-line however, the range of basal activities directed by the LTR clones within each isolate appeared to be greater in isolates which exhibited rapid growth in *in vitro* studies and were derived from individuals

displaying significant disease progression such as CBL20 – 22 (Table 5.1, pg 192) (Shultz *et al.,* 1990), compared to the intraisolate range of LTR activities directed by clones from the less cytopathic members of the CBL series which had been isolated from asymptomatic individuals such as CBL23 and 24 (Figure 5.2B, pg 198). LTR clones from CBL-20, 21, and 22 directed ranges of basal activities that were at least four times greater than the range of basal activities directed by LTR clones from CBL-23 and CBL-24 in this cell-line.

The highest basal activities were directed by LTR clones from CBL-20 and CBL-21, the two most cytopathic and rapidly replicating CBL isolates. The mean basal activities of the five LTR clones from CBL-20 and the five LTR clones from CBL-21 were 13.35 and 10.30 respectively. In contrast, LTR clones from isolate CBL-23, an isolate exhibiting comparatively poorer cytopathogenicity and slower growth (Table 5.1, pg 192), directed the lowest basal activities. The mean basal activity directed by the five LTR clones from this isolate was 6.27, two-fold lower than the mean basal activity directed by the LTRs cloned from CBL-20. The mean basal activity of the five LTR clones from CBL-24, the second isolate derived from an asymptomatic individual, was also lower than the mean basal LTR activities of both CBL-20 and CBL-21.

When tested by one-way ANOVA however, the mean basal activities of the five CBL isolates were not found to be statistically significantly different from each other ($p=0.2157$) (Figure 5.2B, pg 198). Nevertheless, examination of the data does indicate a non-significant trend towards higher basal activities directed by LTRs cloned from isolates displaying rapid growth kinetics.

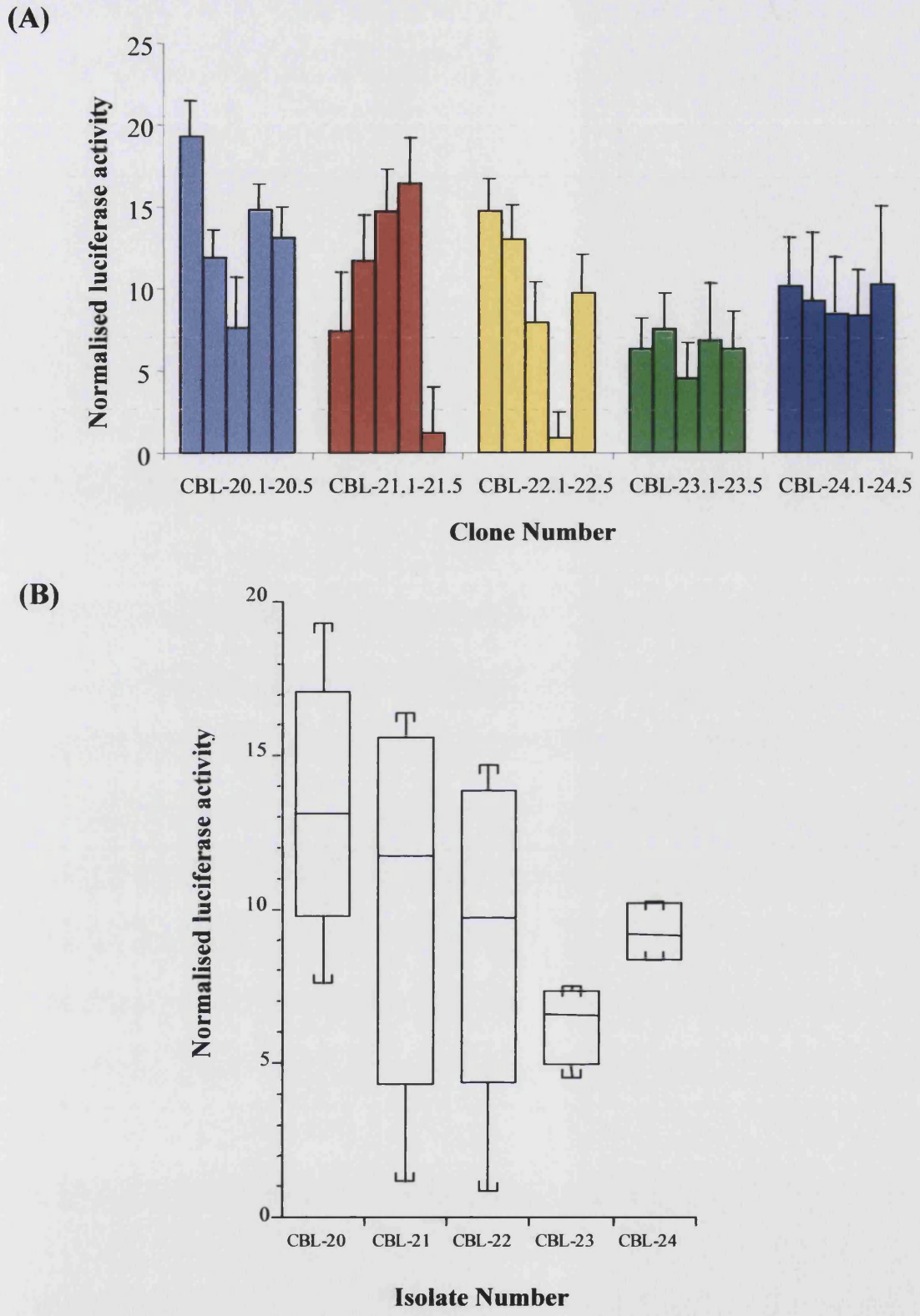


Figure 5.2 (A) Normalised basal activities of 25 cloned LTRs from the CBL series of HIV-2 isolates in Jurkat cells (n=5). Each bar represents the mean normalised basal activity value calculated for each LTR clone from a triplicate transfection, with error bars representing the standard deviation of the three individual values. Cotransfections were performed in 2 x 10^6 Jurkat cells with 1.25μg HIV-2LTR-pGL3E, 0.125μg of pRL-TK and 0.625μg of HSDNA. (B) Box plot representing the range of basal activities directed by five clone for each CBL isolate, the median basal LTR activity for each isolate is marked by a line (n=5). No significant difference in basal LTR activity was noted between the five members of the CBL series p=0.215.

5.3.2.2 Tat-induced activity of HIV-2 LTRs cloned from the CBL series of HIV-2 isolates in the Jurkat cell-line

In order to assess and compare the transcriptional response of the cloned LTRs produced in chapter 4 to the HIV-2 Tat protein (Tat2), a *tat* expression vector was included in the cotransfection mixture for both Jurkat and THP-1 cell lines. The HIV-2 Tat expression vector – pRSVtat contains the first two exons of HIV-2 ROD Tat2, positioned directly downstream from an RSV promoter region (See Appendix II, Figure II.6, pg 353). This vector was a gift from Dr. Browning, University of Michigan, and has been used in several HIV-2 Tat studies.

Comparative transfection studies were performed with saturating concentrations of Tat2. In this respect, the transcriptional responses that were observed were unlikely to be limited by the amount of Tat2 expressed within the cell, however, it must be noted that this protocol may have masked small transcriptional differences between the promoter regions that exist under biological conditions. Section 5.3.2.2 and 5.3.3.2 of this chapter describe the series of experiments undertaken to evaluate and assess the response of cloned HIV-2 LTRs to the HIV-2 Tat protein within both the Jurkat and THP-1 cell-lines.

5.3.2.2a Evaluation of optimal Tat concentration for transactivation of transcription within the Jurkat cell line.

Initially an experiment was undertaken to determine the optimal concentration of pRSV-tat expression vector required to direct the greatest transcriptional response within the Jurkat cell-line. The level of firefly luciferase expression directed by an LTR cloned from HIV-2 ROD following co-transfection with different concentrations of pRSV-tat was assessed and compared.

A constant amount of 3E-ROD vector DNA was cotransfected into freshly seeded Jurkat cells with the internal control reporter pRL-TK at a ratio of 10:1. A range of concentrations of pRSVtat expression vector was added to the set of cotransfection mixtures prior to transfection, ranging from 100ng to 1000ng. The total amount of DNA transfected was maintained at 2μg by supplementation with Herring Sperm DNA, (exact amounts given in the legend of Figure 5.3, pg 201). Normalised basal LTR activity was assessed by including a cotransfection performed in the absence of the pRSVtat vector. Cells were transfected using the Superfect protocol described in section 2.2.4.6a. A triplicate cotransfection was performed for each concentration within the range. Forty-eight hours post transfection cells were harvested by centrifugation, lysed, and the luciferase enzyme activity measured. Normalised absolute Tat-induced LTR activities were calculated for each transfection by dividing the firefly luciferase

expression by the *Renilla* luciferase expression measured from each cell lysate. A mean value for each triplicate cotransfection was then calculated and divided by the normalised basal LTR activity obtained from the co-transfection performed in the absence of the pRSVtat vector in order to determine the fold increase in LTR activity above basal levels, or *trans*activation response, observed at the different pRSVtat concentrations. Plotting the *trans*activation response directed at the different concentrations of pRSVtat expression vector revealed that levels of Tat *trans*activation increased with increasing levels of Tat2 up to a concentration of 600ng (Figure 5.3, pg 201). At this concentration, the expressed Tat2 protein induced a 100-fold increase in LTR activity above the basal level. A decrease in Tat *trans*activation was observed at concentrations greater than 600ng. For subsequent assessment of the *trans*activation response of LTR clones to the HIV-2 Tat protein by transfection of Jurkat cells, 600ng of pRSVtat expression vector were included in the co-transfection mixtures.

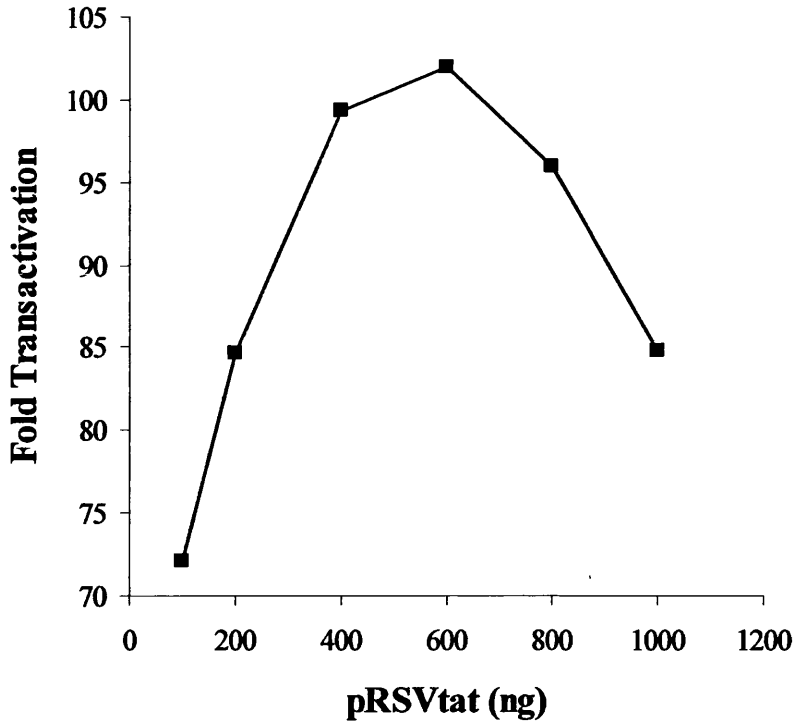


Figure 5.3 Relative *trans*activation of the HIV-2 LTR mediated by various levels of Tat2 in Jurkat cells. A constant level (1μg) of 3E-ROD (pGL3E containing the HIV-2 ROD LTR) was co-transfected with 100ng of pRL-TK and increasing levels (100-1000 ng) of pRSVtat (HIV-2 Tat2). The total amount of DNA transfected per culture was maintained at 2μg by supplementation with HSDNA. Firefly and *Renilla* luciferase activity was measured 48h post transfection and normalised absolute Tat-induced LTR activities were calculated for each transfection by dividing firefly luciferase expression by *Renilla* luciferase expression measured from each cell lysate. The fold increase in promoter activity observed at the different concentrations of pRSVtat expression vector was calculated by dividing the normalised absolute Tat-induced activity obtained in each co-transfection with the normalised basal LTR activity obtained from the co-transfection performed in the absence of the pRSVtat expression vector.

5.3.2.2b Comparison of Tat-induced LTR activities directed by LTRs cloned from the CBL series of HIV-2 isolates

The absolute Tat-induced activities of the LTRs cloned from HIV-2 isolates CBL-20 to CBL-24 were assessed using the dual luciferase assay system (see section 2.2.5). The twenty-five LTR clones representing the five CBL isolates were co-transfected into Jurkat cells with the internal control vector pRL-TK at a ratio of 10:1 and 600ng of the Tat2 expression vector pRSVtat (precise DNA concentrations are given in the legend of Figure 5.4A, pg 204). Co-transfections were performed in triplicate in six well dishes as described in Section 5.3.2.1. Normalised absolute Tat-induced LTR activities were calculated for each transfection lysate by dividing firefly luciferase expression by *Renilla* luciferase expression. Results are presented as the mean normalised absolute Tat-induced activity of each LTR clone, calculated from the three normalised absolute Tat-induced activities of the triplicate transfection (Figure 5.4A, pg

204), and as a box plot representing the range and median of the five clonal absolute Tat-induced LTR activities measured for each CBL isolate (Figure 5.4B, pg 204). All values have had background subtracted. Differences between normalised absolute Tat-induced LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant. In addition, normalised absolute Tat-induced (Figure 5.4A, pg 204) and basal (Figure 5.2B, pg 204) activity values were used to calculate the Tat responsiveness of each LTR clone by dividing the normalised absolute Tat-induced activity by the normalised basal activity for each LTR clone. Results are presented as the mean fold increase in LTR activity above basal levels for each CBL isolate in response to Tat2, with each value representing the mean Tat response of five LTR clones from each isolate (Figure 5.4C, pg 205).

The results from this experiment demonstrated that the activities of each of the twenty-five LTR clones, representing CBL-20 to CBL24, increased considerably over their basal levels when cotransfected with the HIV-2 Tat expression vector in Jurkat cells (Figure 5.2A and 5.4B, pg 204). Significant variation was observed between the absolute Tat-induced activities of LTRs cloned from the same CBL isolate and between LTRs cloned from different CBL isolates (Figure 5.4A, pg 204). However, similar to findings at the basal level of activity, less intra-isolate variation was observed between the Tat-induced activities of the LTRs cloned from isolates CBL-23 and CBL-24, the two HIV-2 isolates taken from asymptomatic individuals (Figure 5.4B, pg 204).

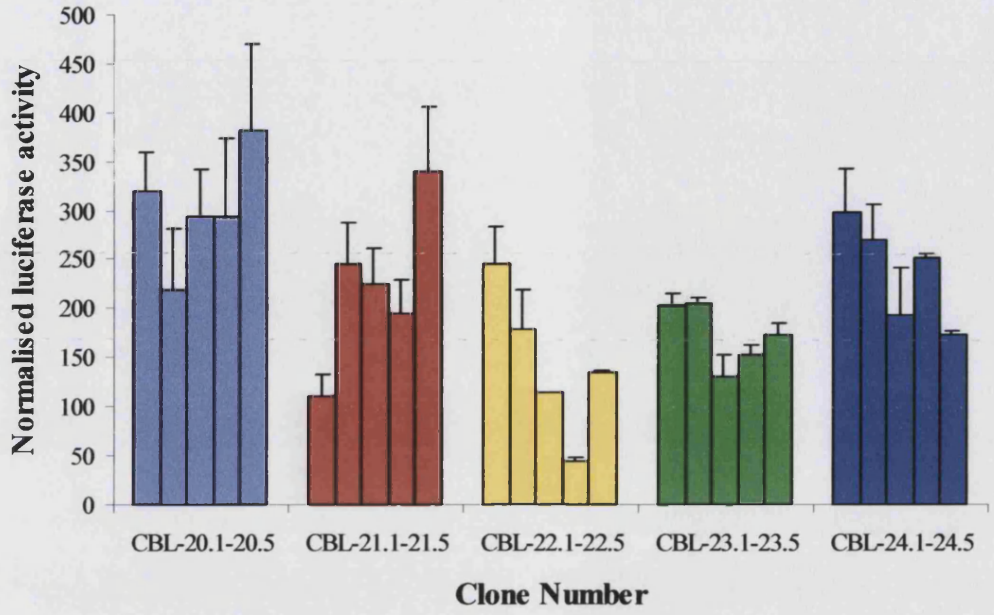
In addition to directing the highest basal activities, the LTRs cloned from CBL-20 were also found to direct the highest absolute Tat-induced activities (Figure 5.4A, pg 204) in this cell-line. The mean normalised Tat-induced activity of the five clones from this isolate was 301.11. In a difference to the basal activities of the LTRs in this cell-line however, LTRs cloned from isolate CBL-22, and not CBL-23, directed the lowest Tat-induced activities in response to the Tat protein. The mean normalised Tat-induced activity of the five clones from CBL-22 was 143.66, two-fold lower than the mean Tat-induced activity directed by the LTR clones from CBL-20. Overall, normalised absolute Tat-induced activities ranged from 44.77 directed by CBL-22 clone 22.4; to 381.54 directed by CBL-20 clone 20.5 (Figure 5.4A, pg 204). The two clones that had displayed extremely low levels of basal activity within the Jurkat cell-line, 22.4 and 21.5 (Figure 5.2A, pg 204), were seen to direct significantly higher levels of transcription in the presence of the Tat protein, particularly clone 21.5 (Figure 5.4A, pg 204). However, clone 22.4 was still the least active of all the LTRs analysed in response to Tat.

When tested by one-way ANOVA the mean absolute Tat-induced LTR activities of at least one of the five CBL isolates were found to be significantly different from the mean absolute Tat-induced LTR activities of the other CBL isolates ($p=0.011$). Multiple comparisons by Tukey-Kramer tests revealed that the Tat-induced activities of the LTRs from CBL-20 were not only higher but statistically higher than the activities of the LTRs from CBL-22 (T-K test $P=0.008$) and CBL-23 (T-K test $P=0.0373$), two strains exhibiting comparatively less cytopathogenicity and slower growth respectively (Table 5.1, pg 192).

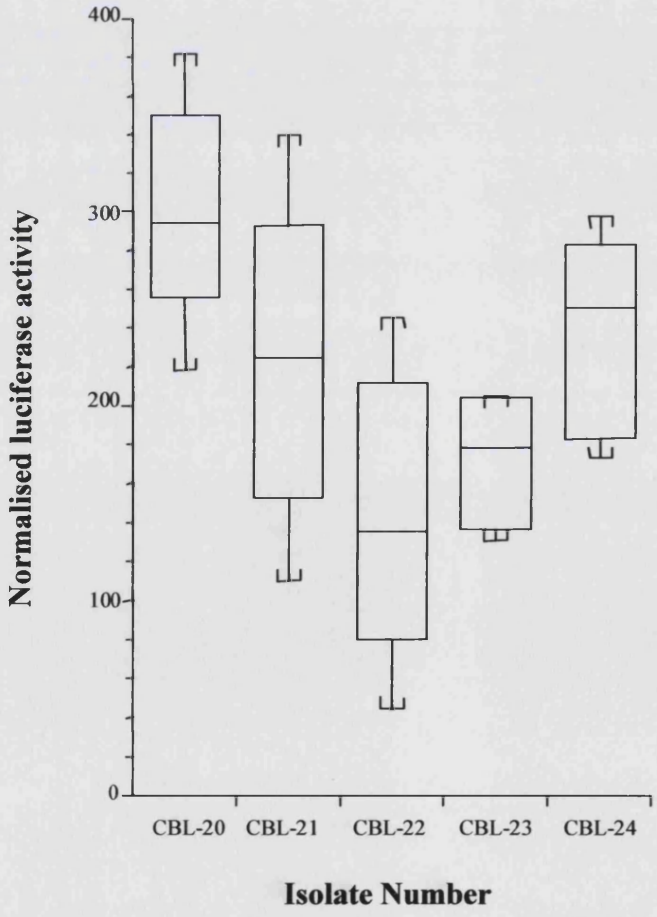
Examination of the absolute Tat-induced activities of the twenty-five LTR clones has revealed a similar trend to that observed at the basal level within the Jurkat cell-line (Figure 5.2A, pg 198, and Figure 5.4A, pg 204). LTRs from isolates exhibiting the most rapid rates of replication i.e. CBL-20 and CBL-21, tended to have higher absolute Tat-induced activities than the LTRs from isolates exhibiting comparatively poorer growth kinetics and cytopathogenicity i.e.CBL-22 and CBL23. However, it must be noted that, on average, the absolute Tat-induced activities of the LTRs cloned from CBL-24, the least cytopathic and slowest growing of the CBL isolates analysed, were higher than the absolute levels of Tat-induced transcription directed by LTRs from CBL-22 and CBL-23, isolates exhibiting more rapid growth kinetics (Figure 5.4A and 5.4B, pg 204) (Table 5.1, pg 192).

Analyses of the mean Tat responsiveness or fold increase in activity above basal levels directed by the CBL LTR clones to the Tat protein, revealed some variation between the five isolates, transcriptional responses to Tat ranged from 18-fold (clones from CBL-21), to 27-fold (clones from CBL-23) (Figure 5.4C, pg 205). LTRs that showed the highest mean fold increase in activity above basal levels in response to the Tat protein however, did not necessarily direct the highest absolute levels of Tat-induced activity. For example, LTRs cloned from CBL-23 were the second least active promoters at the Tat-induced level (Figure 5.4B, pg 204) when compared to LTRs from the other four CBL isolates, and yet directed the greatest mean fold increase in transcription above basal levels in response to Tat (Figure 5.4C, pg 205).

(A)


Normalised luciferase activity
500
450
400
350
300
250
200
150
100
50
0
CBL-20.1-20.5
CBL-21.1-21.5
CBL-22.1-22.5
CBL-23.1-23.5
CBL-24.1-24.5
Clone Number

(B)


Normalised luciferase activity
400
300
200
100
0
CBL-20
CBL-21
CBL-22
CBL-23
CBL-24
Isolate Number

(C)

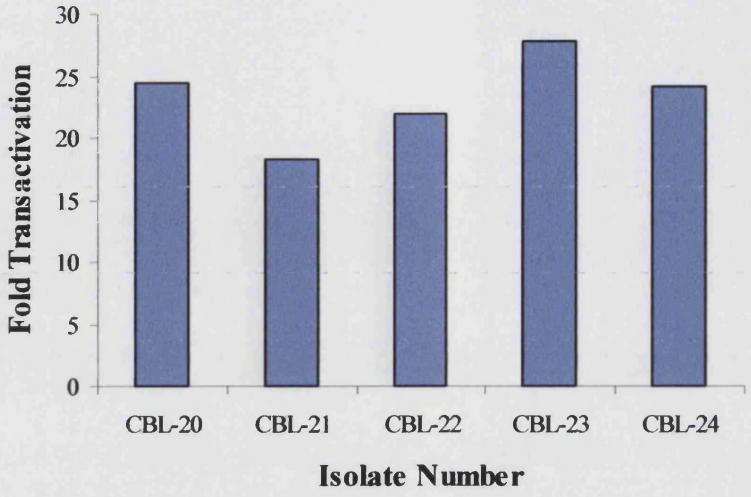


Figure 5.4 (A) Normalised absolute Tat-induced activities of 25 cloned LTRs from the CBL series of HIV-2 isolates in Jurkat cells (n=5). Each bar represents the mean absolute Tat-induced activity value calculated for each LTR clone from a triplicate transfection, with error bars representing the standard deviation of the three individual values. Cotransfections were performed in 2 x 10^6 Jurkat cells with 1.25µg HIV-2LTR-pGL3E, 0.125µg of pRL-TK, 0.6µg of RSV-Tat and 0.025µg of HSDNA. (B) Box plot representing the range of absolute Tat-induced activities directed by five clones for each CBL isolate, the median Tat-induced LTR activity for each isolate is marked by a line (n=5). Analysis by one-way ANOVA revealed significant differences between the mean absolute Tat-induced LTR activities of the five isolates (p=0.011). (C) Mean Tat responsiveness of the LTRs cloned from each of the five CBL isolates (n=5). Tat responsiveness is measured as the fold increase in transcription following Tat transactivation and is calculated by dividing the mean normalised absolute Tat-induced activity of the five LTR clones from each isolate by the mean normalised basal activity of the five LTR clones from each isolate.

5.3.3 Functional analysis of LTRs cloned from five HIV-2 isolates within the THP-1 cell line

HIV-2 LTR functional studies were next performed in the monocyte-like THP-1 cell line; using the same LTR clones as used in the studies with Jurkat cells. LTR function at both the basal and Tat-induced levels was assessed using the same methodologies.

5.3.3.1 Basal activity of HIV-2 LTRs cloned from the CBL series of HIV-2 isolates in the THP-1 cell-line

The basal activities of the twenty-five LTR clones from the five CBL isolates were assessed in the monocyte-like THP-1 cell-line using the dual luciferase assay system (see section 2.2.5). Each clone was cotransfected in triplicate with the internal control plasmid pRL-TK at a ratio of 10:1 according to the Effectene transfection reagent protocol described in section 2.2.4.6b (precise DNA concentrations are given in the legend of Figure 5.5A, pg 208). Luciferase enzyme activity was measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) (see section 2.2.5). Firefly and *Renilla* luciferase background activities were assessed as described in section 2.2.5.2, and subtracted from all corresponding measurements. These cotransfections were performed in parallel to those described in section 5.3.3.2. Results are presented as the mean normalised basal activity for each LTR clone, calculated from the three normalised basal activity values of the triplicate transfection (Figure 5.5A, pg 208), and as a box plot representing the range of the five clonal basal LTR activity values measured for each CBL isolate (Figure 5.5B, pg 208). Differences between LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant.

As seen in the Jurkat cell-line, some level of basal activity was detected from each of the twenty-five LTR clones representing the five members of the CBL series of HIV-2 isolates (CBL-20 – CBL24). However, cell-type-specific transcriptional differences were noted between the basal activities of the LTR clones in the THP-1 cell-line when compared to the basal activities of the same LTR clones when transfected into the Jurkat cell-line. In particular, whilst most active when transfected into Jurkat cells, the LTRs from CBL-20 and CBL-21 directed activities that were not significantly higher than the LTRs from the remaining three isolates when transfected into THP-1 cells. The highest mean basal LTR activity was calculated from the activities of five LTRs cloned from CBL-24, the isolate exhibiting the slowest growth kinetics of the five CBL isolates analysed (Table 5.1, pg 192), and not CBL-20 the most rapidly replicating isolate as observed in Jurkat cells (Figure 5.5B, pg 208). The mean basal LTR activity for

CBL-24 in THP-1 cells was 0.58, five-fold higher than the lowest median basal LTR activity in THP-1 cells, directed by the LTRs derived from isolate CBL-23 (Figure 5.5B, pg 208). The LTR clones from CBL-23 had also directed the lowest basal activities when transfected into Jurkat cells (Figure 5.2A and B, pg 198).

In general, as observed in Jurkat cells, the basal activities of the LTR clones in THP-1 cells were found to be variable within and between the different CBL isolates (Figure 5.5A and B, pg 208). Overall, the activities of the twenty-five clones ranged from 0.02 to 1.31. However, again as observed in Jurkat cells, very little variation was observed between the basal activities of the LTRs cloned from CBL-23, and, with the exception of clone 24.4, between the remaining four clones of CBL-24, the two isolates derived from asymptomatic individuals (Figure 5.5B, pg 208).

In contrast to the trend observed in the basal and absolute Tat-induced activities of the thirty LTR clones when transfected into the Jurkat cell-line, LTR clones from isolates displaying rapid growth kinetics (CBL-20, CBL-21) did not tend to be and more active at the basal level than the LTRs derived from the slower growing CBL isolates (CBL-23, CBL-24) when transfected into the THP-1 cell-line. When tested by one-way ANOVA, there was no overall statistical difference between the mean basal LTR activities of the five CBL isolates in the THP-1 cell-line (P=0.259).

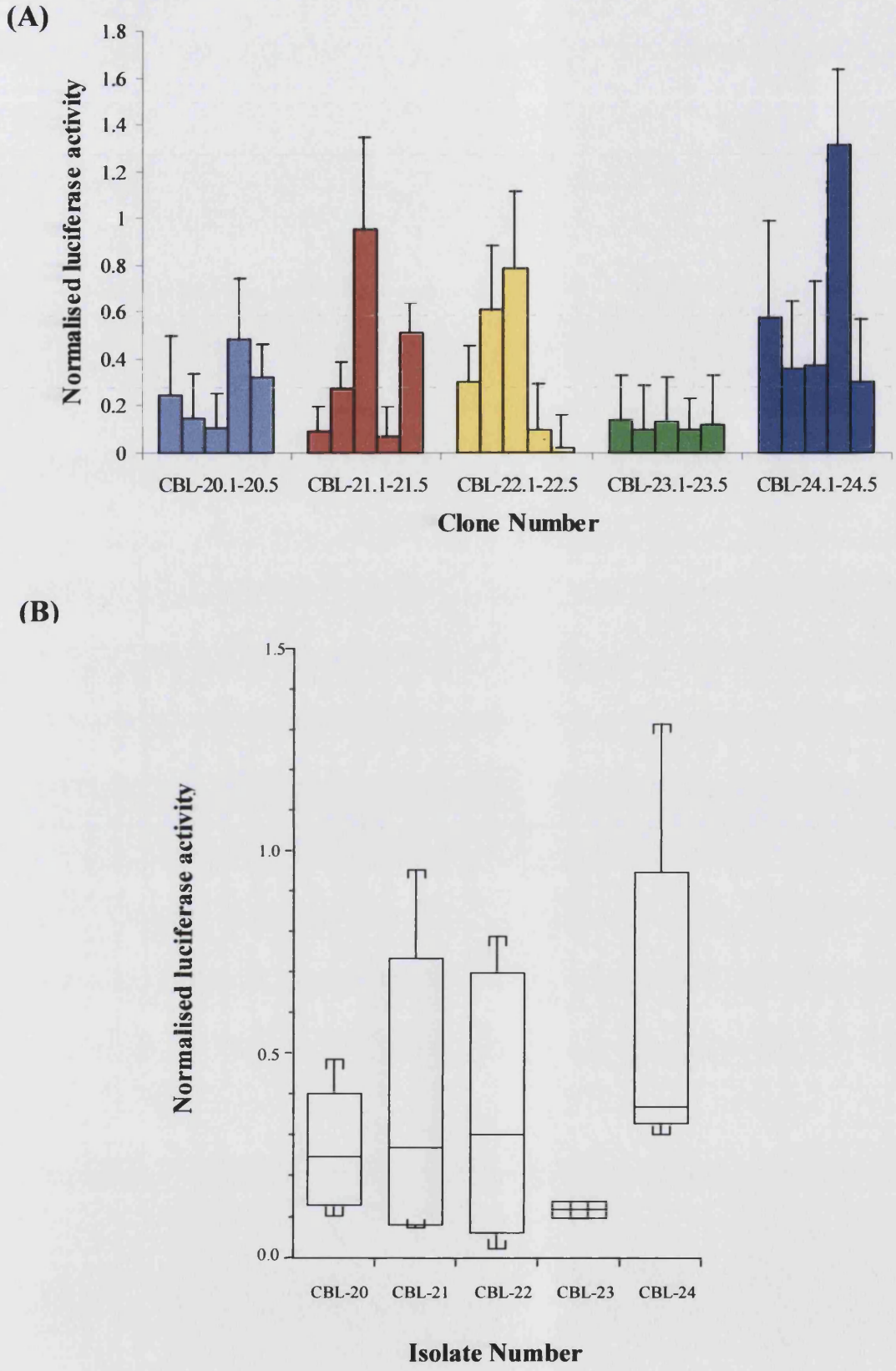


Figure 5.5 (A) Normalised basal activities of 25 cloned LTRs from the CBL series of HIV-2 isolates in THP-1 cells (n=5). Each bar represents the mean normalised basal activity value calculated for each LTR clone from a triplicate transfection, with error bars representing the standard deviation of the three individual values. Cotransfections were performed in 2 x 10^6 THP-1 cells with 0.6μg HIV-2LTR-pGL3E, 0.06μg of pRL-TK and 0.14μg of HSDNA. (B) Box plot representing the range of basal activities directed by five clone for each CBL isolate, the median basal LTR activity for each isolate is marked by a line (n=5). No significant difference in basal LTR activity was noted between the five members of the CBL series (p=0.259).

5.3.3.2 Tat-induced activity of HIV-2 LTRs cloned from the CBL series of HIV-2 isolates within the THP-1 cell-line

5.3.3.2a Evaluation of optimal Tat concentration for transactivation of transcription within the THP-1 cell line.

Before the Tat-induced activities of the twenty-five LTR clones could be assessed in the THP-1 cell-line, the optimal concentration of pRSV-tat expression vector required to direct the greatest transcriptional response in this cell-line was initially determined.

A constant amount of 3E-ROD vector DNA (pGL3E containing the HIV-2 $pROD_{10}$ LTR) was co-transfected into freshly seeded THP-1 cells with the internal control reporter pRL-TK at a ratio of 10:1. A range of concentrations of pRSVtat expression vector was added to a set of co-transfection mixtures prior to transfection, ranging from 5ng to 500ng. Precise DNA concentrations within the transfection mixtures are given in the legend of Figure 5.6, pg 210. Basal LTR activity was assessed by including a co-transfection performed in the absence of the pRSVtat expression vector as described in section 5.3.3.1. The total amount of DNA transfected per 2 x 10^6 THP-1 cells was maintained at 0.8μg by supplementation with Herring Sperm DNA. At forty-eight hours post transfection cells were harvested by centrifugation, lysed and the luciferase activity measured. Normalised absolute Tat-induced LTR activities were calculated for each transfection and divided by the normalised basal LTR activity obtained from the co-transfection performed in the absence of the pRSVtat expression vector in order to determine the fold increase in LTR activity above basal levels, or *trans*activation response.

Plotting the *trans*activation response directed at the different concentrations of pRSVtat expression vector revealed that levels of Tat *trans*activation increased with increasing levels of pRSV Tat up to and including a concentration of 125ng (Figure 5.6, pg 210). At this concentration, the expressed Tat2 protein induced a 40-fold increase in LTR activity above basal levels. A decrease in *trans*activation was observed at concentrations greater than this. For subsequent assessment of the transactivation response of LTR clones to the HIV-2 Tat protein by transfection of THP-1 cells, 125ng of pRSVtat expression vector was included in the co-transfection mixture.

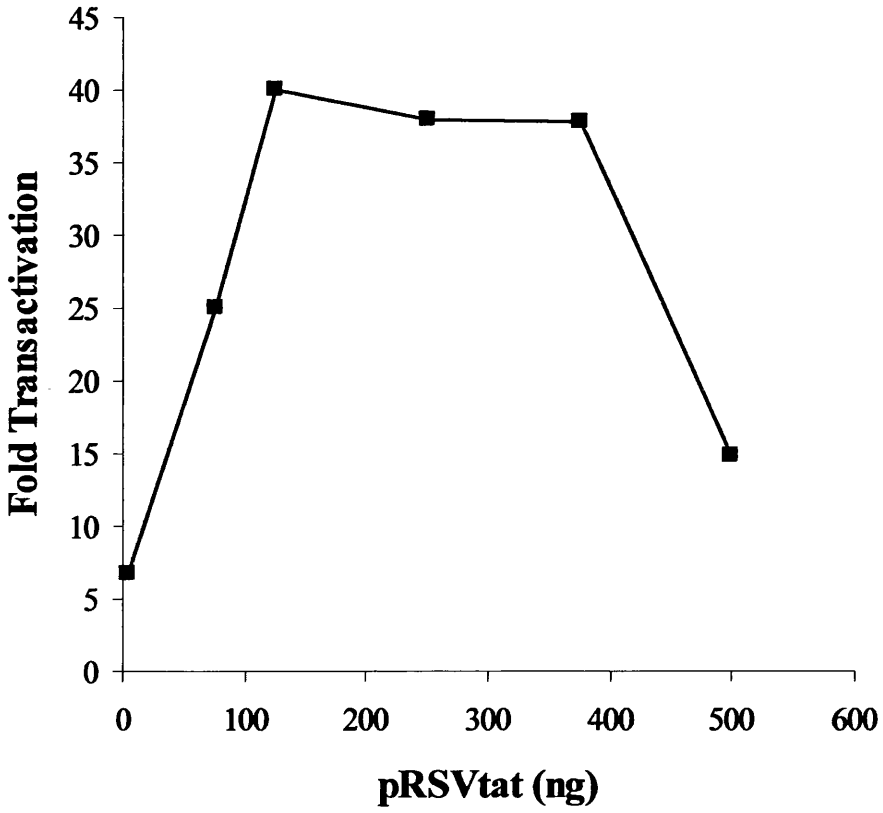


Figure 5.6 Relative *trans*activation of the HIV-2 LTR mediated by various levels of Tat2 in THP-1 cells. A constant level (300ng) of 3E-ROD (pGL3E containing the HIV-2 ROD LTR) was co-transfected with 30ng of pRL-TK and increasing levels (5-500 ng) of pRSVtat (HIV-2 Tat2). The total amount of DNA transfected per culture was maintained at 0.8μg by supplementation with HSDNA. Firefly and *Renilla* luciferase activity was measured 48h post transfection and normalised absolute Tat-induced LTR activities were calculated for each transfection by dividing firefly luciferase expression by *Renilla* luciferase expression measured from each cell lysate. The fold increase in promoter activity observed at the different concentrations of pRSVtat expression vector was calculated by dividing the normalised absolute Tat-induced activity obtained in each co-transfection with the normalised basal LTR activity obtained from the co-transfection performed in the absence of the pRSVtat expression vector.

5.3.3.2b Comparison of Tat-induced LTR activities directed by LTRs cloned from the CBL series of HIV-2 isolates within the THP-1 cell-line

The absolute Tat-induced activities of the twenty-five LTR clones were assessed in the THP-1 cell-line using the dual luciferase assay system (see section 2.2.5). Each LTR clone was cotransfected, in triplicate, with the internal control vector pRL-TK at a ratio of 10:1 and 125ng of the Tat2 expression vector pRSVtat (precise DNA concentrations are given in the legend of Figure 5.7, pg 214). Normalised absolute Tat-induced LTR activities were calculated for each transfection lysate by dividing firefly luciferase expression by *Renilla* luciferase expression. Results are presented in the same format used for the Jurkat data set. Differences between normalised absolute Tat-induced LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant.

As observed in Jurkat cells, the activity of each of the twenty-five LTR clones, representing CBL-20 to CBL24, was raised considerably over basal levels when cotransfected with the HIV-2 Tat expression vector into THP-1 cells (Figure 5.5A, pg 208, and Figure 5.7A, pg 213). Only LTRs within clones CBL-22.5 and CBL-23.4 appeared to direct particularly low absolute Tat-induced activities in this cell-line. Both clones had also directed low basal activities in this cell-line (Figure 5.5A, pg 208). Despite their low Tat-induced activities however, analysis of the responsiveness of each of these LTRs to the Tat protein revealed increases in activity above basal levels in excess of 17-fold.

Further comparison of the basal and Tat data obtained in THP-1 cells (Figure5.5A, pg 208, and Figure 5.7A, pg 213) revealed however, that an LTR with a particularly low basal activity did not necessarily direct a particularly low Tat-induced activity. In particular clones 20.3, 21.4 and 22.4 had all shown comparatively low basal activities in THP-1 cells (Figure 5.5A, pg 208) and yet directed amongst the highest activities in response to the Tat protein within the same cell-line (Figure 5.7A, pg 213). Similarly, while some LTR clones directed both high basal and Tat-induced activities (such as 20.4 and 21.3), others that had been highly active at the basal level in THP-1 cells, such as 22.3 and 24.4, were only of average or below average activity at the Tat-induced level in the same cell-line. Comparable findings were also noted in the Jurkat data set.

Significant variation was observed between the absolute Tat-induced activities of the LTRs cloned from the same CBL isolate and LTRs cloned from different isolates within the CBL series when the LTR clones were transfected into THP-1 cells (Figure 5.7A, pg 213). Activities ranged from 1.53 directed by

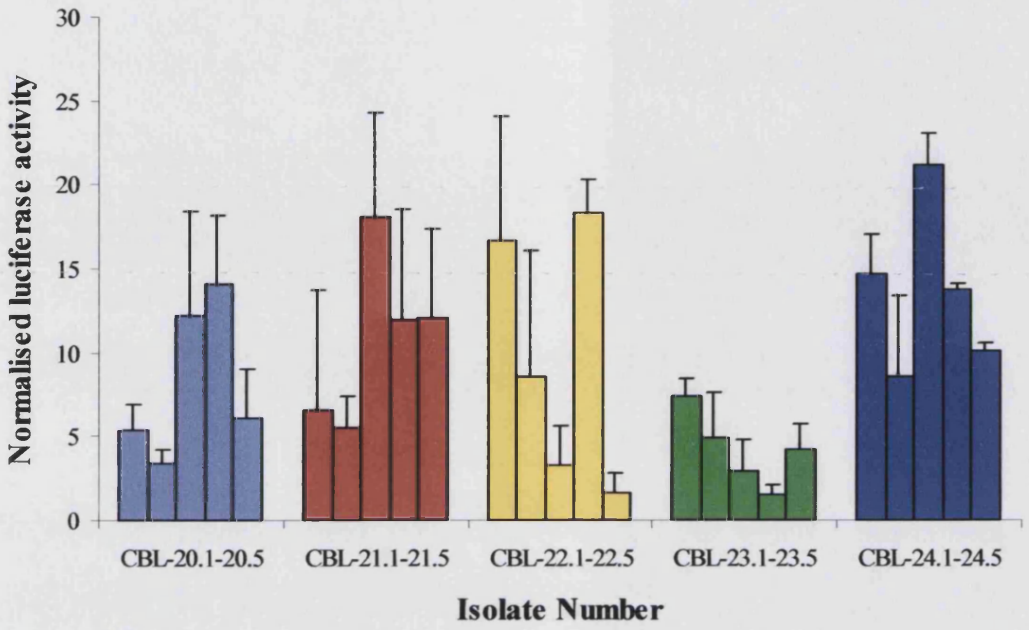
CBL-23 clone 23.4, to 21.11 directed by CBL-24 clone 24.3 (Figure 5.7A, pg 213). In this cell-line, as in the Jurkat cell-line, the lowest intra-patient variation was observed between the activities of LTRs cloned from CBL-23. The range of activities displayed by the clones from this isolate was two-fold lower than the intra-isolate range of LTR activities from CBL-22, the isolate displaying the greatest range of LTR activities in this cell-line.

LTRs cloned from CBL-24, the slowest growing of the CBL isolates, and not CBL-20 as observed in Jurkat cells, directed the highest mean absolute Tat-induced LTR activity within the THP-1 cell-line, while, in addition to displaying the narrowest range of activities, LTRs cloned from CBL-23, an isolate exhibiting more rapid growth kinetics than CBL-24, directed the lowest mean absolute Tat-induced LTR activity. The mean Tat-induced LTR activity of CBL-24 was 13.6, compared to the mean Tat-induced LTR activity of CBL-23, 4.2.

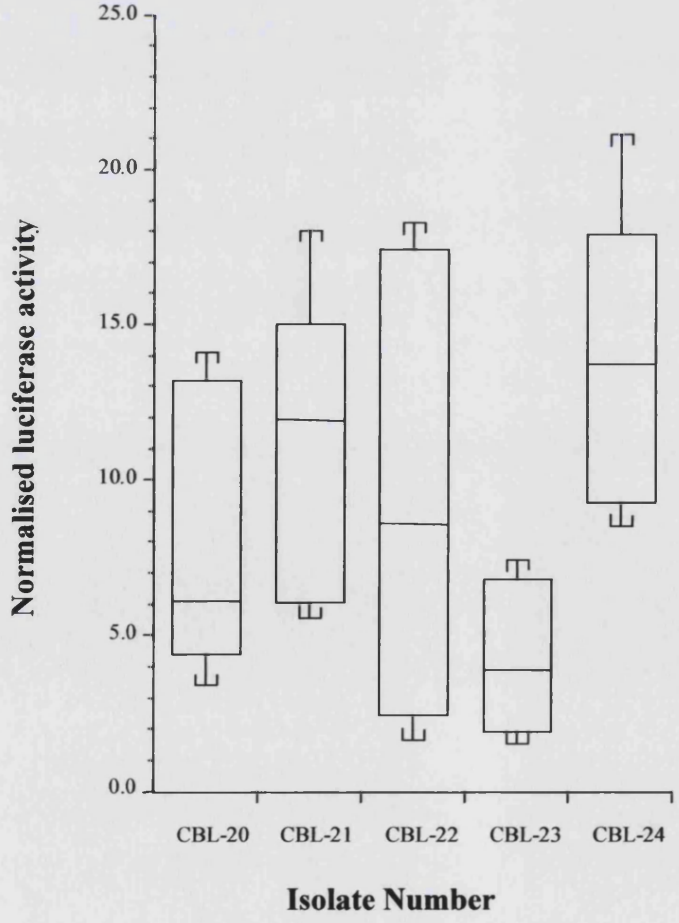
Similar to the basal activity results for these clones when transfected into THP-1 cells, but in contrast to their activities when transfected into Jurkat cells, when tested by one-way ANOVA, the mean absolute Tat-induced LTR activities of the five CBL isolates in the THP-1 cell-line were not found to be statistically significantly different from each other ($p=0.149$). When transfected into THP-1 cells LTRs cloned from CBL isolates of rapid growth characteristics did not tend to direct higher activities than LTRs cloned from slower growing CBL isolates.

Analyses of the mean Tat responsiveness of the promoters to the Tat protein in THP-1 cells revealed some variation between the LTRs from the different CBL isolates. Transcriptional responses to Tat ranged from 23-fold (clones from CBL-24), to 35-fold (clones from CBL-23) (Figure 5.7C, pg 214). Similar to the Jurkat data set, LTRs that had directed the highest median absolute Tat-induced LTR activity, did not necessarily direct the greatest transcriptional response above basal levels in the presence of the Tat protein in THP-1 cells.

(A)


Normalised luciferase activity
30
25
20
15
10
5
0
CBL-20.1-20.5
CBL-21.1-21.5
CBL-22.1-22.5
CBL-23.1-23.5
CBL-24.1-24.5
Isolate Number

(B)


Normalised luciferase activity
25.0
20.0
15.0
10.0
5.0
0.0
CBL-20
CBL-21
CBL-22
CBL-23
CBL-24
Isolate Number

(C)

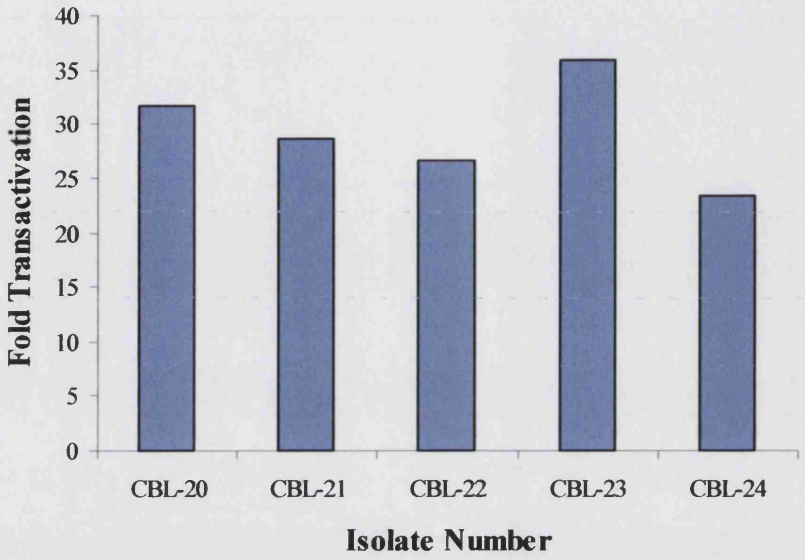


Figure 5.7 (A) Normalised absolute Tat-induced activities of 25 cloned LTRs from the CBL series of HIV-2 isolates in THP-1 cells (n=5). Each bar represents the mean absolute Tat-induced activity value calculated for each LTR clone from a triplicate transfection, with error bars representing the standard deviation of the three individual values. Cotransfections were performed in 2 x 10^6 THP-1 cells with 0.6μg HIV-2LTR-pGL3E, 0.06μg of pRL-TK, 0.125μg of RSV-Tat and 0.015μg of HSDNA (B) Box plot representing the range of absolute Tat-induced activities directed by five clone for each CBL isolate, the median Tat-induced LTR activity for each isolate is marked by a line (n=5). Analysis by one-way ANOVA revealed no significant differences between the mean absolute Tat-induced LTR activities of the five isolates (p=0.149). (C) Mean Tat responsiveness of the LTRs cloned from each of the five CBL isolates (n=5). Tat responsiveness is measured as the fold increase in transcription following Tat transactivation and is calculated by dividing the mean normalised absolute Tat-induced activity of the five LTR clones from each isolate by the mean normalised basal activity of the five LTR clones from each isolate.

5.4 Summary and discussion

The basal and Tat-induced activities of LTRs directly cloned from the DNA of a series of HIV-2 isolates displaying different growth characteristics were compared in two biologically relevant cell-lines in order to examine the importance of promoter activity with respect to replicative capacity. Previous characterisation of the biological and molecular variability of the HIV-2 CBL isolates had suggested that, as with HIV-1, a relationship existed between the *in vitro* cytopathogenicity of the different isolates and the clinical disease status of the patient from whom they were originally isolated (Shultz *et al.*, 1990). Therefore, comparative analysis of the LTRs from each of these isolates also examined the importance of promoter activity with respect to disease progression.

In order to determine the practicable number of LTR clones required to represent each CBL isolate within comparative transfection studies, the variation in basal activity directed by multiple LTRs cloned from a single CBL isolate, CBL-22, was initially examined. Using the optimised dual-luciferase assay system significant functional variation was demonstrated between the basal LTR activities of several of the ten LTRs cloned directly from CBL-22 isolate DNA ($p<0.0001$), with activities falling over a 34-fold range. The basal activity of LTR clone 22.6 was found to be statistically higher than all other LTRs cloned from this isolate, in addition, the LTR activity of clone 22.2 was found to be statistically higher than the LTR activities of clones 22.9, and 22.10. The differences in activities of the LTRs derived from CBL-22 likely reflect the existence of functionally significant polymorphisms within the core transcription factor binding elements of the LTR in the CBL-22 quasispecies, and also suggests that despite tissue culture adaptation and numerous passages in culture, functional variability persists within the quasispecies of the CBL series of HIV-2 isolates. Based on this result, and experimental design considerations it was concluded that multiple LTR clones would give a more representative view of promoter activity within the viral quasispecies of an isolate than a single LTR clone. Five LTR clones were chosen as a practicable number to represent each of the CBL isolates in the subsequent comparative transfection experiments.

Using the optimised dual-luciferase assay system a detectable level of basal activity was measured from each of twenty-five LTR clones, representing five members of the HIV-2 series of CBL isolates, in both the Jurkat and THP-1 cell-lines. Ten of the LTRs analysed were cloned from CBL-23 and CBL-24, the two isolates in the analyses whom were derived from asymptomatic individuals (Shultz *et al.*, 1990). These data suggest therefore, that the asymptomatic state of the patients from whom isolates CBL-23 and

CBL-24 were obtained was not determined by the existence of completely defective LTRs within the circulating quasispecies of these patients.

When tested by one-way ANOVA there was no overall statistical difference between the mean basal LTR activities of the five CBL isolates in either the Jurkat or THP-1 cell-line ($p=0.2157$, and $p=0.259$ respectively). Examination of the basal data obtained within the Jurkat cell-line did indicate however, a non-significant trend towards higher and more varied basal activities being directed by the LTRs cloned from isolates that exhibited rapid growth kinetics and high cytotoxicity in several cell-lines (Shultz *et al.*, 1990) and were taken from patients displaying significant disease progression. The highest basal activities within the Jurkat cell-line were directed by the LTRs cloned from isolates CBL-20, 21, and 22. Moreover, the ranges of basal activity directed by the LTRs from each of these three isolates were on average four times greater than the range of basal activities directed by LTRs cloned from CBL-23 and CBL-24. In contrast, the lowest basal activities in Jurkat cells were directed by the LTRs cloned from CBL-23 and CBL-24; both isolates displayed comparatively slower growth kinetics than CBL-20 and CBL-21, were less cytopathogenic, and were obtained from patients who were asymptomatic at the time of virus isolation (Shultz *et al.*, 1990).

Examination of the basal activities of the same LTRs when transfected into the THP-1 cell-line, did not reveal any similar trends in activity. In contrast to the Jurkat data, the LTRs cloned from CBL-24 (the isolate that displayed the slowest growth kinetics *in vitro*), directed the highest mean basal activities in THP-1 cells. The observation that several of the LTR clones had directed the highest activities in one cell-line but were not necessarily the most active LTRs when transfected into a second cell-line, is in agreement with previously published results analysing HIV-1 LTR activity (Corboy *et al.*, 1992, Moses *et al.*, 1994, Estable *et al.* 1996, Krebs *et al.*, 1998, McAllister *et al.*, 2000), and indicates that some *in vivo*-selected HIV-2 LTRs may confer transcription and replication advantages in specific cell types. Only LTRs cloned from CBL-23 and LTR clone 22.4 were consistently found to direct low basal activities in both THP-1 and Jurkat cells, perhaps suggesting functionally significant sequence differences between the core transcription factor binding elements of these LTRs compared to the other LTRs in the analyses.

When evaluating the optimal Tat2 expression vector concentration for *trans*activation of transcription within both Jurkat and THP-1 cells optimal *trans*activation was achieved when 600ng of pRSVTat was cotransfected into Jurkat cells, and 125ng when cotransfected into THP-1 cells. Significant decreases in *trans*activation were noted in concentrations exceeding the given amounts in both cell-lines. Similar

findings have been reported by Fenrick *et al* while investigating the relative *trans*activation of the HIV-1 and HIV-2 LTRs by various levels of Tat1 and Tat2 (Fenrick *et al.,* 1989). The observed decrease in *trans*activation is thought to result from the toxicity of the Tat protein itself to the transfected cell.

At the Tat-induced level of activity each of the twenty-five LTR clones, representing CBL-20 to CBL24, directed a significant increase in activity above basal levels in both Jurkat and THP-1 cells when cotransfected with the Tat2 expression vector, a further indication that the LTRs cloned from each of the CBL isolates were fully functional. Within both cell-lines, variation was observed between the mean Tat responsiveness of the LTRs from each of the CBL isolates analysed, with Tat responsiveness being defined as the fold increase in activity above the basal activity of that LTR. On average, levels of transcription were found to increase between 25 and 40–fold above basal levels in the presence of Tat2. Interestingly, LTRs that directed the greatest fold increase in activity above their basal activities in response to Tat were not necessarily those LTRs that directed the highest absolute Tat-induced activities. For example, the LTRs from CBL-23 were on average the most Tat-responsive LTRs in both cell-lines however; they directed the lowest and second lowest absolute Tat-induced activities when transfected into THP-1 and Jurkat cells respectively. Therefore, the replicative capacities of the HIV-2 isolates appear to relate more to the actual magnitude of promoter activity than the *responsiveness* of the promoter to proteins such as Tat2.

Similar to results at the basal level of activity in Jurkat cells, LTRs cloned from the isolates exhibiting the most rapid rates of replication i.e. CBL-20 and CBL-21, tended to direct higher absolute Tat-induced activities within Jurkat cells than the activities of the LTRs from isolates exhibiting comparatively poorer growth kinetics and cytopathogenicity i.e. CBL-23. On average, the Tat-induced activities of the LTRs from CBL-20 were two-fold higher than the activities of CBL-22 and CBL23. Moreover, when tested statistically the mean Tat-induced activity of the LTRs from CBL-20 in Jurkat cells was found to be statistically significantly higher than the mean Tat-induced activity of the LTRs from CBL-22 (T-K test P=0.008) and CBL-23 (T-K test P=0.0373). LTRs cloned from CBL-20 had also directed the highest levels of basal activity within Jurkat cells, although statistical analysis had shown that the activities of the LTRs from this isolate were not statistically higher than the activities of the LTRs from the other viral isolates.

CBL-20 represents a markedly different *in vitro* characterised isolate phenotype compared to the phenotypes exhibited by CBL-22 and CBL-23. CBL-20 is highly cytopathic, exhibits rapid growth, and

was derived from an AIDS patient, while strains of CBL-22 and CBL-23 exhibit poorer cytopathogenicity, slower growth, and were derived from an ARC and an asymptomatic patient respectively (Shultz *et al.,* 1990). It has been suggested that differences in the basal and inducible activities of LTRs from HIV-1 viral isolates may contribute to the variability in the replicative capacity of HIV strains, with higher activities providing replicative advantage (Golub *et al.,* 1990, Englund *et al.,* 1991, Zhang *et al.,* 1997a, Berkhout *et al.,* 1999). In a study performed by Golub *et al,* a 3-fold difference between the basal activities of the LTRs from two intrapatient HIV-1 isolates was found to be responsible for a three-fold difference in replication between the two HIV-1 isolates (Golub *et al.,* 1990). Moreover, Anderson and co-workers demonstrated that the 1.6 to 2-fold higher Tat-induced activity of the LTR from infectious SIV_{mac} molecular clone $SIV_{mac}239$ when compared to the LTR of $SIV_{mac}251$ paralleled the higher levels of virus and viral RNA found in macaques infected with this clone compared to those infected with $SIV_{mac}251$ (Anderson & Clements, 1992). Monkeys infected with $SIV_{mac}239$ had persistent viremia and usually followed an acute disease course, whereas $SIV_{mac}251$-infected animals had transient viremia, little virus in tissue and followed a more prolonged chronic disease course. Therefore, it is possible that the statistically significant 2-fold difference observed between the Tat-induced activities of the LTRs cloned from CBL-20, when compared to CBL-22 and CBL-23, may in part determine the different replicative capacities of the three isolates. Similarly, the observed differences may in turn reflect differences in the overall promoter activity of the likely phenotypes circulating within the patients from whom the isolates were derived, and the difference in the progression status of the three patients from whom the isolates were obtained.

Nevertheless, while directing higher activities, the differences observed between the activities of the LTRs from CBL-21, also a rapidly replicating viral strain, and the activities of LTRs from the remaining four isolates, and between LTRs from CBL-20 and CBL24, the viral strains exhibiting the most extreme differences in growth kinetics and cytopathogenicity, all within the Jurkat cell-line, were not found to be statistically significant ($P=0.5227$).

In addition, when transfected into the THP-1 cell-line the LTRs cloned from the rapidly replicating isolates were not any more active at the Tat-induced level than the LTRs cloned from the isolates displaying slower growth kinetics. LTRs cloned from CBL-24 (the least cytopathic and slowest growing of the CBL isolates), and not CBL-20 (the most rapidly growing and cytopathic CBL isolate), directed the highest mean absolute Tat-induced activity. In contrast to the statistical differences observed between the mean Tat-induced LTR activities of CBL-20, CBL22, and CBL23 within the Jurkat data set, when tested

by one-way ANOVA, the mean Tat-induced LTR activities of the five CBL isolates were statistically indistinguishable from each other in THP-1 cells.

In summary, significant differences have been noted between the mean basal and Tat-induced LTR activities of the CBL series when transfected into the T-cell-like Jurkat cell-line. In general, the LTRs cloned from the cytopathic, rapidly replicating isolates isolated from patients with either AIDS or ARC tended to have higher activities than the LTRs cloned from slower growing, less cytopathic isolates isolated from asymptomatic individuals. When the same LTR clones were transfected into the monocyte-like THP-1 cell-line however, the mean basal and Tat-induced LTR activities of the members of the CBL series were found to be statistically indistinguishable from each other. In THP-1 cells, LTRs cloned from the cytopathic, rapidly replicating isolates were not any more likely to have higher activities than the LTRs cloned from slower growing and less cytopathic isolates. In fact, LTRs derived from the least cytopathic and slowest growing isolate directed the highest levels of basal and Tat-induced activities in this cell-line.

The observation of such cell-type-specific transcriptional differences and of the cell-type-specific trends towards higher basal and Tat-induced activities directed by the LTRs from the rapidly replicating isolates in Jurkat cells, raises the possibility that some HIV-2 LTRs could confer transcriptional and replicative advantage within specific cell-types. In this instance when transfected into a T-cell-like cell-line, LTRs cloned from rapidly replicating isolates derived from patients displaying significant disease progression appear to have transcriptional advantage over LTRs cloned from isolates exhibiting slower growth kinetics that were derived from asymptomatic patients. Interestingly, the range of basal and Tat-induced activities directed by the LTRs cloned from CBL-20, 21, and 22 were found to be larger in the THP-1 and more particularly the Jurkat cell-line, than the range of basal and Tat-induced activities directed by LTR clones taken from CBL-23 and CBL-24.

Cell-type-specific transcriptional differences between the activities of HIV-1 LTRs have been demonstrated extensively (Corboy *et al.*, 1992, Moses *et al.*, 1994, Estable *et al.* 1996, Krebs *et al.*, 1998, Jeeninga *et al.*, 2000), as has the tissue-compartmentalised phylogenetic clustering of HIV-1 LTR variants (Ait-Khaled *et al.*, 1995, Corboy & Garl, 1997), leading to the suggestion that the LTR is involved in cell type-specific gene expression and viral replication (Ait-Khaled *et al.*, 1995, Corboy & Garl, 1997, Krebs *et al.*, 1998, Jeeninga *et al.*, 2000). More recently, evidence has been presented to suggest that in this role the HIV LTR is also a potential contributor to viral tropism (Corboy & Garl,

1997, McAllister *et al.,* 2000). The identification by Clark *et al* (Clark *et al.,* 1995) of a binding site (peri-κB) within the HIV-2 LTR that was required for transcription within monocytic but not T cell-lines has indicated that the HIV-2 LTR could also contribute to cell type-specific gene regulation, viral replication, and viral tropism. It is interesting to note that the activities of the LTRs from CBL-20 and CBL-24 in particular, appear to reflect the tropism of the viral isolates from whom they were derived. LTRs from CBL-20, the isolate that grew rapidly in primary culture (reverse transcriptase activity – 6000cpm in 5 days, Schultz *et al.,* 1990) and T cell-like cell-lines such as CEM, H9, and MOLT4, directed the highest activities when transfected into the T-cell-like Jurkat cell-line. Whereas, LTRs from CBL-24, the isolate that grew slowly in primary culture (reverse transcriptase activity – 6000cpm in 28 days, Schultz *et al.,* 1990) and could only be grown in the T cell-like MOLT4 cell-line after several attempts directed some of the lowest activities in the Jurkat cell-line. In contrast, the LTRs from this isolate directed the highest activities when transfected into the monocyte-like THP-1 cell-line. The significance of these findings to HIV-2 pathogenesis requires further investigation. Nevertheless, since *in vitro* culturing of HIV places viral isolates within highly selective environments the conclusions that can be drawn from this data set are limited and must be made with caution. However, the findings appear to indicate that the HIV-2 LTR maybe contributing to cell-type specific gene expression and perhaps viral tropism. The next chapter in this thesis describes the functional analysis of LTRs cloned from the uncultured PBMCs of HIV-2 infected patients displaying either long-term non-progression, or rapid progression to symptomatic disease.

The results from chapter 5 have demonstrated that the optimised dual luciferase assay system is capable of reproducibly detecting statistically significant differences between the activities of HIV-2 LTRs derived from different DNA sources in different cell-lines. This has enabled reliable comparative analyses of HIV-2 LTR function to be achieved.

Chapter 6.

Functional Characterisation of Naturally Occurring HIV-2 LTR Genotypes from LTNP and RP Gambian Patients.

6.1 Introduction

Unlike HIV-1 few naturally occurring HIV-2 LTR genotypes have been functionally characterised (Tong-Starksen *et al.*, 1990, Arya, 1991, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al* 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995). In order to investigate further any correlation between LTR activity and disease progression rates in HIV-2 infected individuals, and to extend the findings presented in Chapter 5, the activity of LTRs cloned directly from the uncultured PBMCs of HIV-2 patients with non-progressing disease and patients showing progressive disease were assessed and compared. Patients with relatively rapid progressing HIV-2 AIDS, with disease courses similar to HIV-1 infection, have been observed both in The Gambia (Ariyoshi *et al.*, 2000) and among West Africans resident in Europe (van der Ende *et al.*, 1996). However, rapid progression is generally uncommon amongst HIV-2 infections. Therefore, comparing the function of LTRs from these unusual patients with LTR function in more typical slow progressors may reveal key differences in promoter activity.

We have examined the function of LTRs cloned from the uncultured PBMCs of three LTNP and three RP HIV-2 infected patients (described in Chapter 3) from The Gambia. These individuals were identified as HIV-2 seropositive at the MRC hospital, Fajara, and were recruited into a longitudinal study for a period of clinical follow-up extending at least 56 months. The basal and Tat-induced activities of multiple LTR clones from each of the six patients have been measured in two biologically relevant cell-lines, the Jurkat T-cell-line and the monocytic THP-1 cell-line, using the optimised dual luciferase assay system (described in Chapter 4).

6.2 Summary of patient details

PBMC samples were obtained from six HIV-2 seropositive patients attending the MRC hospital in Fajara, The Gambia (procedures were performed by colleagues at the MRC laboratories, Fajara, The Gambia). Each of the six patients had been recruited to a longitudinal study in which patient samples were taken at six monthly intervals (Ariyoshi *et al.*, 2000). Disease progression rates were determined based on clinical data, decline of CD4% over the period of follow-up (determined by colleagues at the MRC laboratories, Fajara, The Gambia), and plasma viral load measurements (determined by Dr. N. Berry, NIBSC, London). Rapid progressor (RP) patients were defined as those with CD4 declining from greater than 20% to less than 13% within 20 months, and whose viral load measurements exceeded 50,000 copies/ml (RP1-RP3). Long-term non-progressor (LTNP) patients had stable CD4 counts for more than 56 months, and had viral load measurements of <1000 copies/ml (LTNP1-LTNP3). LTRs were cloned from PBMC samples obtained as close to the time of RNA viral load measurements as possible.

MRC Sample code (assigned sample number)	Longitudinal RNA (Load measurements – date and (copies/ml)) (determined by Dr.N.Berry, NIBSC, London)	Date of PBMC sample from which LTRs were cloned	CD4% at time of LTR cloning (determined by colleagues at MRC Laboratories, Fajara, The Gambia)	Disease Status
N72617 (LTNP1)	27/08/91-**undetectable** 13/06/95- **2568**	09/11/93	41	Non-progressor
N71746 (LTNP2)	08/10/91-**undetectable** 23/06/92-**undetectable** 07/06/94- **undetectable**	22/06/93	49	Non-progressor
N15120 (LTNP3)	30/04/92-**undetectable** 27/08/92-**undetectable** 17/05/94 **undetectable**	06/06/96	48	Non-progressor
N72462 (RP1)	13/08/92 - **100,000**	19/10/93	9	Rapid progressor
W1333 (RP2)	18/08/92 - **55,450**	18/08/92	22	Rapid progressor
N74044 (RP3)	06/10/92 - **500,000** 07/01/93 - **99,500** 14/06/94 - **183,000**	14/06/94	13	Rapid progressor

Table 6.1 Origin of the PBMC samples from HIV-2 infected individuals. Multiple LTRs were cloned from each of the six patients and were used to assess LTR function at the basal and Tat-induced level. Undetectable RNA load measurements are defined as those < 200copies/ml.

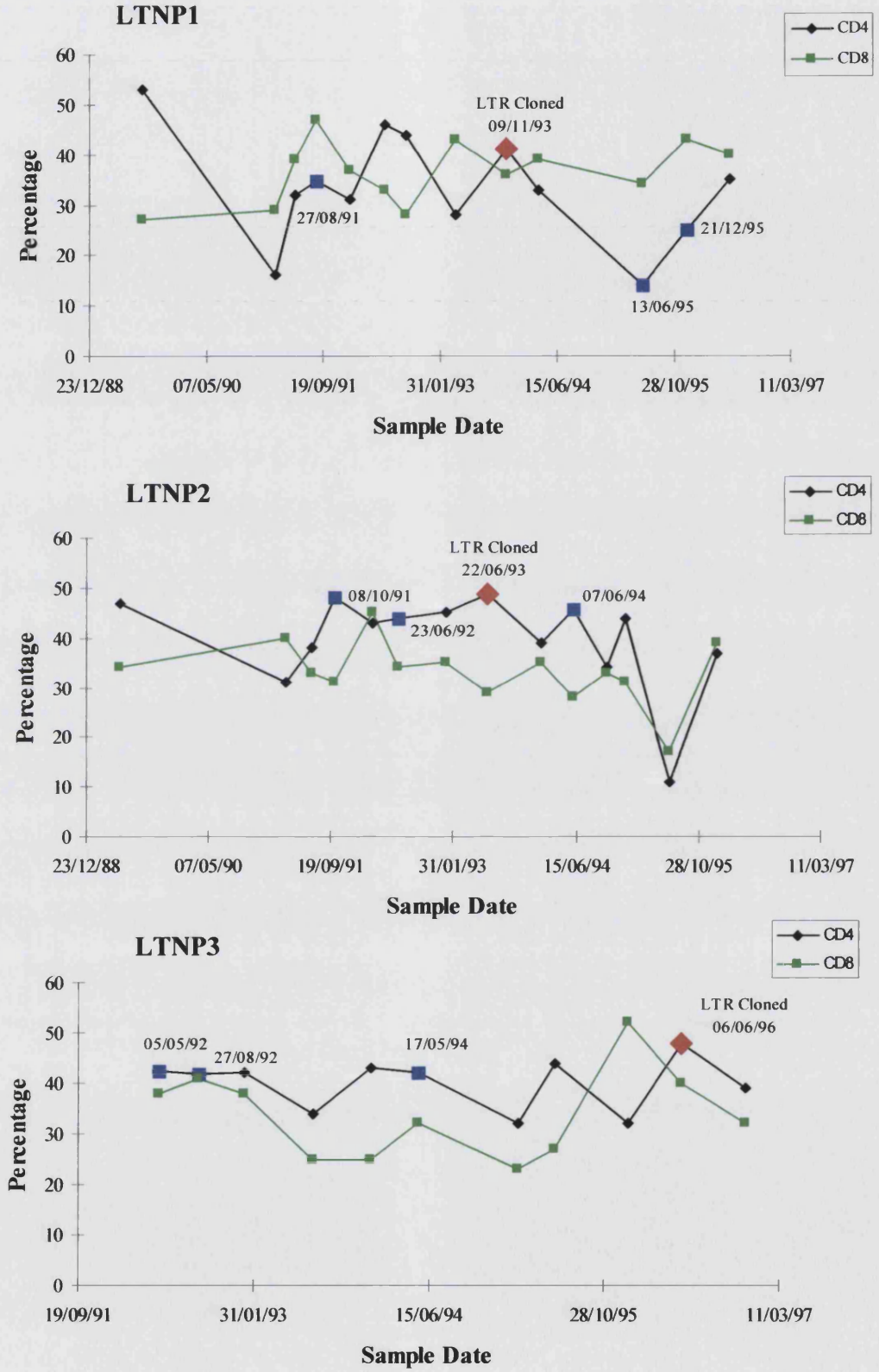


Figure 6.1 Graphical representations of the longitudinal CD4 and CD8 % profiles of the three LTNP patients in the analyses (determined by colleagues at the MRC Laboratories, Fajara, The Gambia). Red diamond symbols represent the date of the PBMC sample from which LTRs were directly cloned. Blue square symbols represent the date at which plasma HIV-2 RNA has been measured.

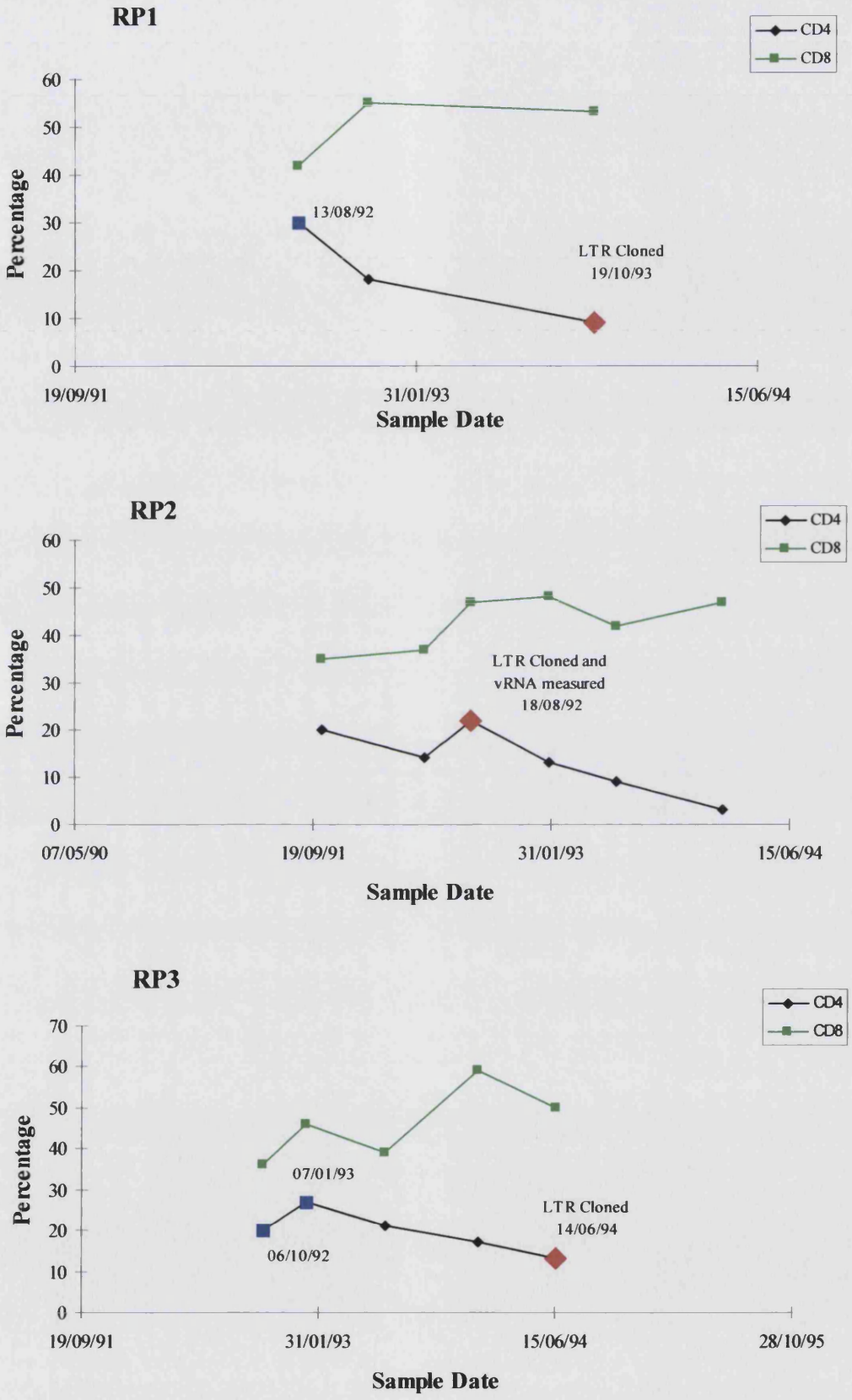


Figure 6.2 Graphical representations of the longitudinal CD4 and CD8 % profiles of the three RP patients in the analyses (determined by colleagues at the MRC Laboratories, Fajara, The Gambia). Red diamond symbols represent the date of the PBMC sample from which LTRs were directly cloned. Blue square symbols represent the date at which plasma HIV-2 RNA has been measured.

6.3 Transcriptional activity of naturally occurring HIV-2 LTR genotypes from RPs and LTNPs

Assessing and comparing the activities of LTRs cloned directly from the uncultured PBMCs of six patients in total examined the role of HIV-2 LTR function in determining disease progression status further. Three patients had been characterised as long-term non-progressors (LTNPs), the remaining three patients were characterised as rapid progressors (RPs) (see Table 6.1, pg 223, and Figures 6.1, pg 224, and 6.2, pg 225).

6.3.1 Evaluation of intrapatient LTR functional variability

In order to enable a judgement to be made about the number of LTR clones required to give a representative view of the LTR activity within each patient in subsequent comparative transfection experiments, the variation in the basal activities of naturally occurring HIV-2 LTR genotypes derived directly from the PBMCs of two HIV-2 infected Gambian patients, one randomly selected from each progression group, was examined.

Sixteen LTR clones (eight from long-term non-progressor 3 - LTNP3, eight from rapid progressor 2 - RP2) produced using the nested LTR PCR – restriction digest cloning procedures described in Chapter 4 were cotransfected into Jurkat cells, in triplicate, with the internal control vector pRL-TK at a ratio of 10:1, using the Superfect transfection reagent protocol described in section 2.2.4.6a. Precise DNA concentrations within the transfection mixtures are given in the legend of Figure 6.3, pg 228. Luciferase activities derived from the HIV-2 promoters are expressed as normalised activities by dividing the firefly luciferase expression by the *Renilla* luciferase expression measured using the dual-luciferase assay system (Promega Ltd) and luminometry (Canberra-Packard Topcount) from each cotransfection. Results are presented as the mean normalised basal activity value for each LTR clone, calculated from the three normalised basal activity values of each triplicate cotransfection. Differences between LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant.

When transfected into Jurkat cells the LTRs within each of the sixteen reporter constructs, eight cloned from LTNP3 and eight cloned from RP2, had a detectable basal activity (Figure 6.3A, pg 228). Basal LTR activity was found to be variable between the eight LTRs cloned from each of the two patients. The

basal activities directed by the LTRs from RP2 varied over a 7-fold range (Figure 6.3A, pg 228) whilst the basal activities of LTRs from LTNP3 varied over a 6-fold range (Figure 6.3B, pg 228).

When tested by one-way ANOVA the differences observed between the mean normalised basal activities of the eight clones derived from RP2 and the eight clones derived from LTNP3 were found to be statistically significant (ANOVA p=0.0024 and p=0.0115 respectively). This analysis demonstrated that at least one of the eight LTR clones in each patient data set directed a mean basal activity that was significantly different from the mean basal LTR activities of the remaining LTR clones from that patient which would not be explained by sampling variation.

Multiple comparisons within the RP2 data set (Figure 6.3A, pg 228) by Tukey-Kramer tests revealed that the mean normalised basal LTR activity of clone RP2.7 was significantly higher than the mean normalised basal activities of six of the seven other RP2 LTR clones (T-K test p=0.02 – 0.002). The activity of the LTR within this clone was at least two-fold and in some cases up to six-fold greater than the activities of the other LTR clones in this data set. RP2.4 was the only LTR clone from RP2 that directed a level of basal activity that, while lower, was not statistically different from the activity of the LTR within clone RP2.7. The basal activity of the LTR within clone RP2.4 was found to be statistically higher than the activities of clones RP2.2 (T-K test p=0.05) and RP2.6 (T-K test p=0.04).

Within the LTNP3 data set (Figure 6.3B, pg 228) the basal activity of LTR clone LTNP3.1 was found to be significantly higher than the basal activities of clones LTNP3.5 (T-K test p=0.02) and LTNP3.8 (T-K test p=0.03). The activity of the LTR within clone LTNP3.1 exceeded the activities of these two clones by five-fold.

Since statistically significant differences were observed between the basal activities of LTRs cloned from the same patient the results indicated the necessity for multiple LTR clones to give a more representative view of naturally occurring promoter activity within the viral quasispecies of each Gambian HIV-2 infected patient than a single LTR clone. Since clones from each of six Gambian patients were to be transfected at least in triplicate into two different cell lines within the subsequent comparative analyses, and were to be assessed for basal and Tat-induced activity within both cell lines, a practicable figure of five clones was chosen to represent the promoter region within the quasispecies of each Gambian patient.

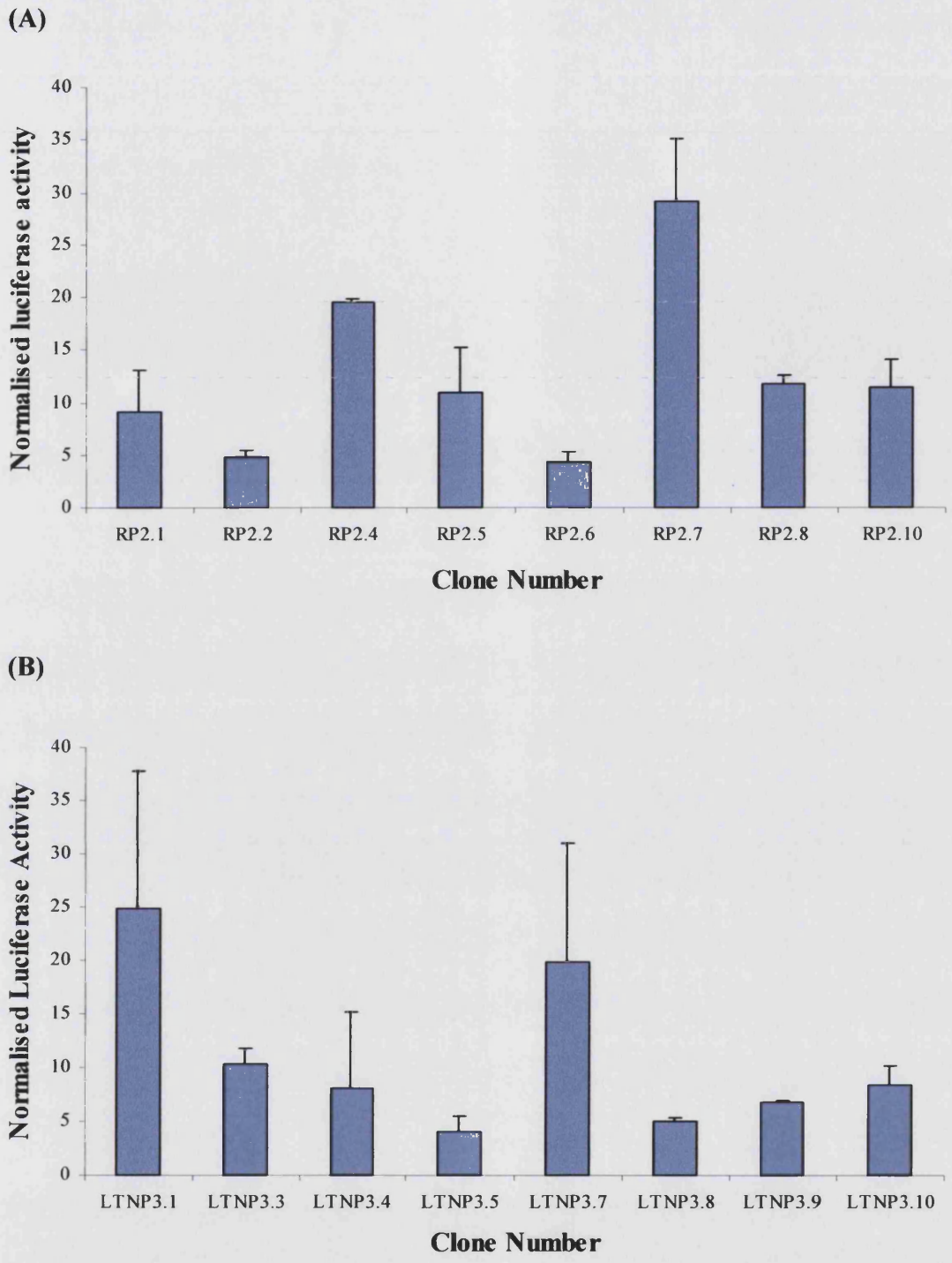


Figure 6.3 (A) Differential normalised basal activity of eight LTR clones from RP2. (B) Differential normalised basal activity of eight LTR clones from LTNP3. The cloned LTRs were tested for basal activity in 2 x 10^6 Jurkat cells. Cotransfections were performed with 1.26μg HIV-2LTR-pGL3E, 0.14μg of pRL-TK and 0.6μg of HSDNA. Background luciferase activity was assessed by performing a transfection with 1.26μg pGL3E and 0.74μg of HSDNA, and was subtracted from all measurements. Each bar represents the mean normalised basal activity value calculated for each LTR clone from a triplicate transfection, with error bars representing the standard deviation of the three individual values. Luciferase activities derived from the HIV promoters were normalised by dividing the firefly luciferase expression by the *Renilla* luciferase expression measured from each cell lysate.

6.3.2 Functional analysis of LTRs cloned from LTNP and RP patient PBMCs within the Jurkat cell line

HIV-2 LTR functional studies were initially undertaken within the T-cell like Jurkat cell line. LTR function at the basal and Tat-induced levels for each of the six HIV-2 infected Gambian patients, three LTNPs and three RPs, was assessed and compared, each patient was represented by five LTR clones produced as described in Chapter 4.

6.3.2.1 Basal activity of HIV-2 LTRs cloned from Gambian cohort patients

The level of basal activity directed by LTRs cloned from six Gambian HIV-2 infected patients was assessed using the optimised dual luciferase assay system (see section 2.2.5). A total of thirty LTR clones, five from each of the six patients (produced by the nested-PCR and restriction digest cloning procedures described in Chapters 3 and 4), were cotransfected in triplicate into Jurkat cells with the *Renilla* luciferase internal control vector pRL-TK at a ratio of 10:1 according to the protocol described in section 2.2.4.6a (precise DNA concentrations are given in the legend of Figure 6.4A, pg 232). Luciferase activities derived from each of the thirty HIV-2 promoters were then normalised with respect to *Renilla* luciferase activity by dividing the firefly luciferase expression by the *Renilla* luciferase expression measured from each cotransfection (see section 2.2.5). Over the course of the study this entire experiment was repeated a total of three times, and in parallel to the Tat activity analyses (see section 6.3.2.2). Results are presented as the mean normalised basal activity value for each LTR clone, calculated from nine normalised basal activity values obtained in three independent triplicate cotransfections (Figure 6.4A, pg 232), and as a box plot representing the range of the five clonal basal LTR activity values measured for each Gambian patient (Figure 6.4B, pg 233). The range of the fifteen clonal basal LTR activities for each progression group and the range of the three mean basal LTR activities calculated for the three patients in each of the two progression groups are also presented as box plots (Figure 6.5A, pg 234, and Figure 6.5B, pg 234, respectively). A line marks the median basal LTR activity for each patient or progression group, the box representing the distance between the first and third quartiles, the leading and trailing 'whiskers' corresponding to the minimum and maximum respectively (Figure 6.4B, pg 233, 6.5A, pg 234, and 6.5B, pg 234). Differences between LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant.

Each of the thirty LTR clones taken from a total of six Gambian patients, three LTNP, three RP, was found to direct a detectable level of basal activity when transfected into Jurkat cells (Figure 6.4A, pg 232). Significant variation was observed between the basal activities of LTRs cloned from the same Gambian patient and between the basal activities of LTRs cloned from different Gambian patients. The greatest intra-patient ranges in basal activity were seen between LTRs cloned from RP2 and RP3 (Figure 6.4A, pg 232, and 6.4B, pg 233), the range of activities directed by the LTRs cloned from RP3 was at least two-fold greater than the intra-patient ranges of all three LTNPs and one RP, RP1.

Overall, LTRs cloned from the RP group of patients directed basal activities ranging from 4.52 to 12.80 with a median basal activity of 7.62 (n=15) (Figure 6.5A, pg 234). The median basal activity of the LTR clones from the LTNP group of patients was lower than that directed by LTRs from the RP group of patients at 5.89 (n=15), with activities ranging from 2.45 to 8.91. One way ANOVA analysis of the basal activities of the thirty LTR clones when grouped as two progression groups (n=30), revealed that within Jurkat cells the mean basal LTR activity of the group of fifteen RP LTR clones was statistically significantly higher (ANOVA p=0.008) than the mean basal LTR activity of the group of fifteen LTNP LTR clones (Figure 6.5A, pg 234). However, when this analysis was performed comparing the three LTNP patient mean basal activity values (calculated from the five clonal LTR activities for each patient) with the three RP patient mean basal activity values, the difference between the mean basal activities of the two progression groups became non-significant (ANOVA p=0.254) (Figure 6.5B, pg 234).

Of the thirty HIV-2 LTR clones tested, those cloned from RP2 and RP3 directed the highest overall levels of basal LTR activity within the Jurkat cell-line. Clones RP3.5, RP3.1, RP2.5 and RP2.3 were the most active LTRs at the basal level, whilst the lowest basal activities within this cell-line were directed by the five LTRs cloned from LTNP1 (Figure 6.4A, pg 232). On average, the basal LTR activities of both RP2 and RP3 were at least three fold higher than the mean basal activity of the five LTR clones from LTNP1.

When the mean basal LTR activities of the six patients were individually tested and compared by one-way ANOVA, the mean basal LTR activity of at least one of the six Gambian patients was found to be statistically significantly different to the mean basal LTR activities of the remaining patients in the analysis (ANOVA $p<0.0001$) (Figure 6.4B, pg 233). This test demonstrated that in at least one instance the difference between the individual patient means was greater than the intrapatient variation in LTR activity and was unlikely therefore to have occurred as a result of sampling variation. Multiple comparisons by Tukey-Kramer tests revealed that with the exception of RP1, the mean basal activity of

LTRs from LTNP1 was significantly lower than the mean basal LTR activities of all other patients, both LTNP and RP, in the analyses (T-K test p=0.0219 - <0.0001). The analyses also demonstrated that the mean basal LTR activities of two of the RP patients, RP2 and RP3, were statistically significantly higher than the activities of LTRs from two of the LTNP patients, LTNP1 (T-K test p<0.0001, p<0.0001) and LTNP2 (T-K test p=0.0096, p=0.0138), in addition to RP1 (T-K test p=0.0018, p=0.0026), which, of the three RP patients, directed the lowest overall levels of basal transcription in this cell-line. LTR clones from LTNP3 directed a mean basal LTR activity (7.02) that, while lower than the mean basal LTR activities of RP2 (9.35) and RP3 (9.20), was not statistically distinguishable from them (T-K test p=0.19, p=0.26).

Examination of these data has indicated a non-significant trend in basal LTR activity within the Jurkat cell-line. LTRs cloned from patients displaying rapid rates of disease progression, high levels of plasma viremia and low CD4 counts, (Rapid progressors) (Table 6.1, pg 223), tended to be more active at the basal level within the Jurkat cell-line than LTRs cloned from patients who, for a prolonged period of time have shown no signs of clinical progression, have high CD4 counts and have no detectable plasma viremia (Long-term non-progressors) (Table 6.1, pg 223).

(A)

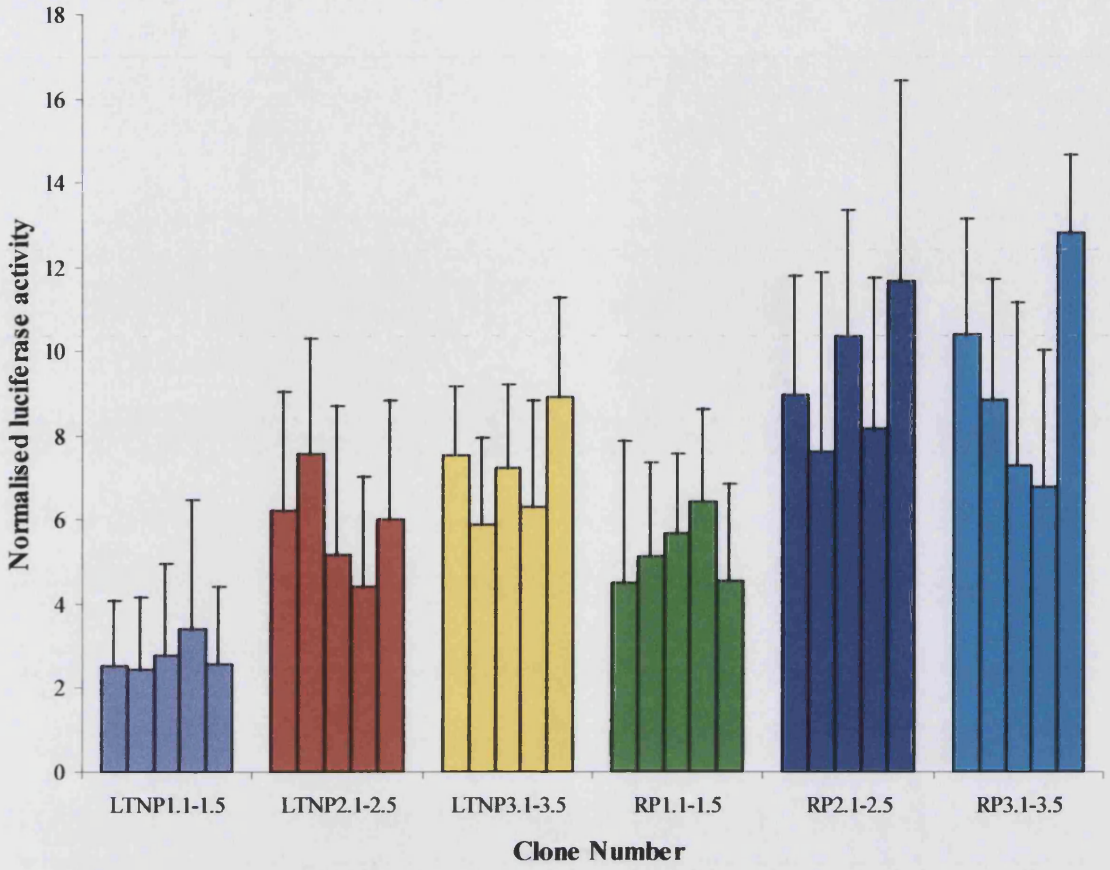


Figure 6.4 (A) Normalised basal activities of 30 cloned LTRs in Jurkat cells (n=6). Five LTR clones represent each of six Gambian HIV-2 seropositive patients, three characterised as LTNPs, three as RPs. Each bar represents the mean normalised basal activity value calculated for each LTR clone from nine normalised basal activity values obtained in three independent triplicate cotransfections, with error bars representing the standard deviation of the nine individual values. Cotransfections were performed in 2 x 10^6 Jurkat cells with 1.25μg HIV-2LTR-pGL3E, 0.125μg of pRL-TK and 0.625μg of HSDNA.

(B)

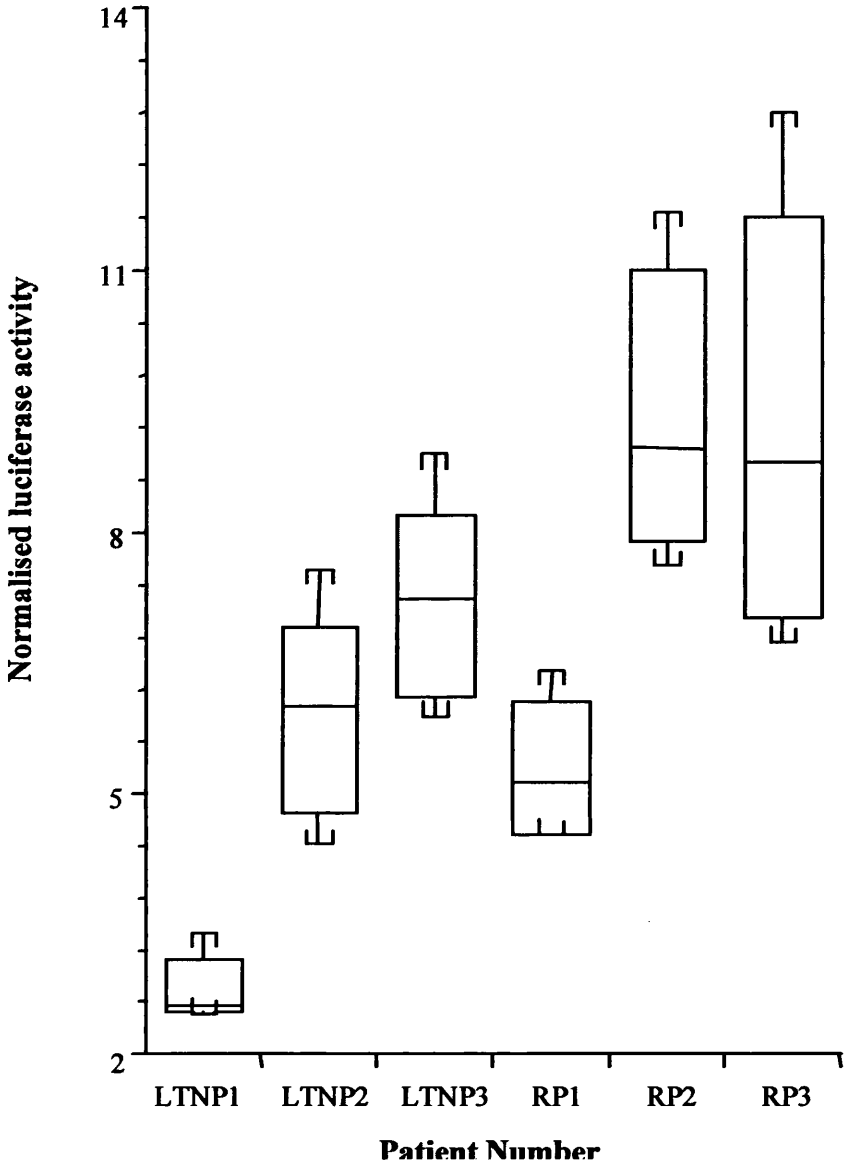
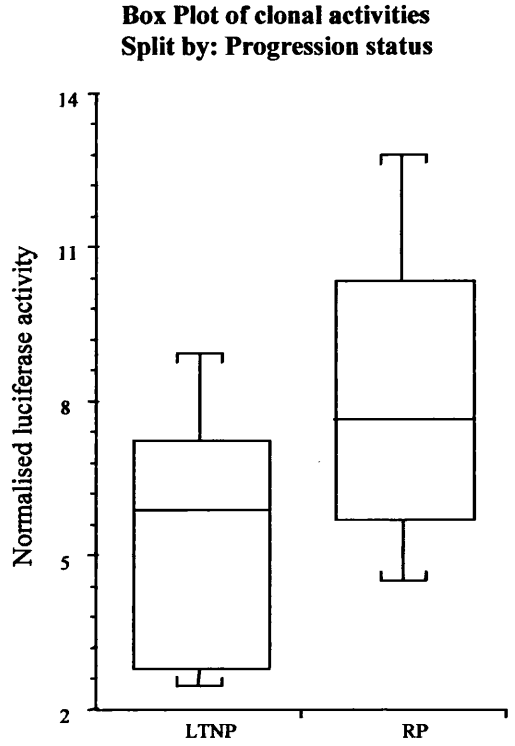


Figure 6.4 (B) Box plot representing the range of basal activities directed by the five LTR clones from each of six patients in Jurkat cells, the median basal LTR activity for each patient is marked by a line (n=6). Testing by one-way ANOVA revealed that the mean basal LTR activity of at least one of the six patients was statistically different from the other patients in the analyses (p=0.0001). Multiple comparisons by Tukey-Kramer tests revealed that the mean basal LTR activity of LTNP1 was, with the exception of RP1, significantly lower than all other patients both LTNP and RP (p=0.0219 – <0.0001). Furthermore, the mean basal LTR activities of RP2 and RP3 were statistically significantly higher than LTNP1 (p<0.0001, p<0.0001) and LTNP2 (p=0.0096, p=0.0138), in addition to RP1 (p=0.0018, p=0.0026).

(A)



(B)

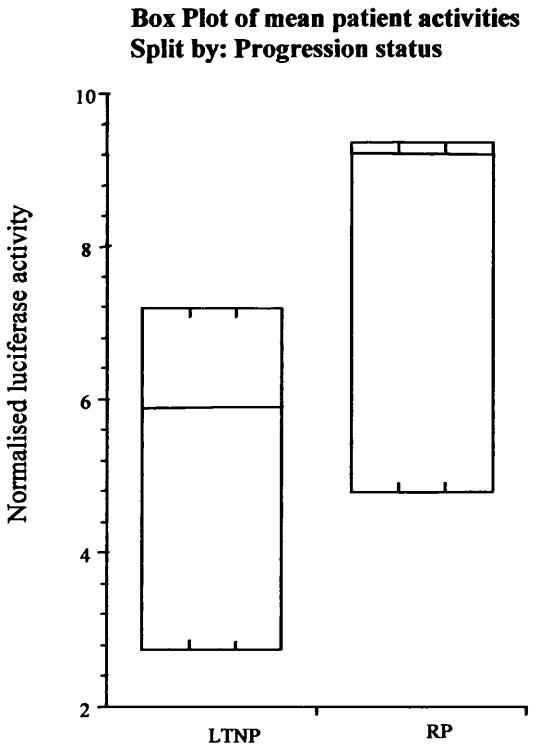


Figure 6.5 (A) Box plot representing the range of basal LTR activities directed by the fifteen clones from each progression group in Jurkat cells (n=6). Each progression group consisted of three patients from whom five LTRs were cloned. Characterisation of patients was based on clinical data available (Table 6.1, pg 223, and Figures 6.1, pg 224, and 6.2, pg 225). Testing by one-way ANOVA revealed that the mean basal LTR activity of the RP group of clones was statistically higher than the mean basal LTR activity of the LTNP group of clones (n=30, p=0.008). (B) Box plot representing the range of the three mean patient basal LTR activities in each progression group (n=6). Statistical comparison of the mean basal LTR activities of the two progression groups using the three patient means from each group (n=6) revealed that the difference between the mean basal LTR activities of the two groups became non-significant (p=0.254).

6.3.2.2 Tat-induced activity of HIV-2 LTRs cloned from the Gambian cohort patients in the Jurkat cell-line

In order to assess and compare the transcriptional response to HIV-2 Tat directed by the HIV-2 LTR genotypes cloned from the three LTNP and three RP Gambian patients, the HIV-2 Tat expression vector RSV-tat was included in the cotransfection mixtures for both Jurkat and THP-1 cell lines.

The thirty LTR clones, five from each of the six Gambian patients were cotransfected into the T-cell like Jurkat cell-line with the *Renilla* luciferase internal control plasmid pRL-TK at a ratio of 10:1 and 600ng of the Tat2 expression vector pRSVtat. Cotransfections were performed in triplicate using the Superfect protocol for the transfection of Jurkat cells (see section 2.2.4.6a). The experiment to determine the basal LTR activity of the Gambian LTR clones, (as described in section 6.3.2.1), was performed in parallel to the Tat experiment described below. Normalised absolute Tat-induced LTR activities derived from each of the HIV-2 promoters were normalised with respect to *Renilla* luciferase activity by dividing the firefly luciferase expression by the *Renilla* luciferase expression measured from each cotransfection (see section 2.2.5). Over the course of the study this entire experiment was repeated a total of three times.

Results are presented as the mean normalised absolute Tat-induced activity for each LTR clone, calculated from nine normalised absolute Tat-induced activity values obtained in the three independent triplicate transfections (Figure 6.6A, pg 239), and as a box plot representing the range and median of the five clonal absolute Tat-induced LTR activity values measured for each patient (Figure 6.6B, pg 240). The range of the fifteen clonal absolute Tat-induced LTR activities for each progression group and the range of the three mean Tat-induced LTR activities calculated for the three patients in each of the two progression groups are also presented as box plots (Figure 6.7A, pg 241, and Figure 6.7B, pg 241, respectively). Differences between normalised absolute Tat-induced LTR activities were compared by analysis of variance (ANOVA) and by the Tukey-Kramer test. Values of $p<0.05$ were considered statistically significant.

In addition to plotting the normalised basal (Figure 6.4A, pg 232) and absolute Tat-induced activities of the thirty LTR clones (Figure 6.6A, pg 239) these values have been used to calculate the Tat responsiveness of each LTR clone by dividing the normalised absolute Tat-induced activity by the normalised basal activity for each LTR clone. Results are presented as the mean fold increase in LTR activity above basal levels directed by each LTR clone in response to Tat2 (Figure 6.8, pg 242).

When cotransfected into Jurkat cells, with the exception of clones from LTNP1, each of the LTRs derived from the remaining five patients directed significantly higher activities in the presence of Tat2 than they had directed at the basal level (Figure 6.4A, pg 232, and 6.6A, pg 239). Strikingly, none of the five LTR clones from LTNP1 directed any increase in activity above basal levels when cotransfected with the Tat2 expression vector into Jurkat cells.

As observed at the basal level of activity within the Jurkat cell-line considerable variation was observed between the normalised absolute Tat-induced activities of LTRs cloned from the same patient and between LTRs cloned from different patients (Figure 6.6A, pg 239). Of the six patients analysed, the greatest intra-patient variation was observed between the activities of the LTRs cloned from each of the three RP patients (Figure 6.6B, pg 240). The intrapatient ranges of LTR activity for RP1, RP2, and RP3 were, in each case, at least two-fold greater than the ranges of activities directed by the five LTRs from each of the three LTNP patients. In addition, the range of Tat-induced activities directed by the five LTR clones from RP3 was found to be 2-fold greater than the range of activities directed by LTRs from the two other RP patients in the analyses, RP1 and RP2.

Overall, LTRs cloned from the RP group of patients directed absolute levels of Tat-induced activity ranging from 180.56 to 665.48 with a median Tat-induced activity of 348.3 (Figure 6.7A, pg 241). The median Tat-induced activity of LTRs from the LTNP group of patients was significantly lower than that directed by LTRs from the RP group of patients at 162.5, with activities ranging from 1.51 to 206.40. On average, RP Tat-induced LTR activities were three times higher than LTNP Tat-induced LTR activities.

With the exception of clone RP1.1, the absolute Tat-induced activities of fourteen of the fifteen LTRs from the RP group of patients were higher than the absolute Tat-induced activities directed by each of the fifteen LTR clones from the three LTNP patients (Figure 6.6A, pg 239). Moreover, the median Tat-induced LTR activities for each of the three RP patients were higher than the median Tat-induced LTR activities for the three LTNP patients within the Jurkat cell-line (Figure 6.6B, pg 240).

One-way ANOVA analysis of the absolute Tat-induced activities of the thirty LTR clones when grouped into two progression groups (n=30), revealed that within Jurkat cells the mean Tat-induced LTR activity of fifteen clones from three RP patients was statistically significantly higher (ANOVA $p<0.0001$) than the mean Tat-induced LTR activity of fifteen clones from three LTNP patients (Figure 6.7A, pg 241). However, as observed at the basal level of activity, when this analysis was performed comparing the three RP patient mean Tat-induced activity values (calculated from the five clonal values for each patient) with

the three LTNP patient mean Tat-induced activity values, the difference between the mean Tat-induced activities of the two progression groups became non-significant (ANOVA p=0.0595) (Figure 6.7B, pg 241).

Of the thirty LTR clones tested the highest absolute normalised Tat-induced activities were directed by LTRs from RP3 (Figure 6.6A, pg 239). Particularly high levels of activity were directed by LTRs clones RP3.3, RP3.1 and RP3.2. Of the LTRs that did produce a *trans*activation response to the Tat2 protein LTRs cloned from LTNP2 directed the lowest absolute Tat-induced activities. On average, the absolute Tat-induced activities of LTRs cloned from this patient were three-fold lower than the activities directed by the LTRs from RP3.

Statistical analysis of the mean absolute Tat-induced LTR activities of the six Gambian patients when not grouped as two progression groups revealed that at least one of the patient means was statistically significantly different to the mean absolute Tat-induced LTR activities of the other patients in the analysis (ANOVA n=30, p<0.0001). Multiple comparisons between the six patient means by Tukey-Kramer tests revealed the mean absolute Tat-induced activity of LTRs from LTNP1 to be significantly lower than the mean Tat-induced LTR activities of the remaining five Gambian patients (T-K test p<0.0001 – p=0.0045). In addition, the mean absolute Tat-induced activity of LTRs from RP3 was significantly higher than the mean Tat-induced LTR activities of all five Gambian patients (T-K test p=0.0032 - <0.0001), including RP1 (T-K test p=0.0002) and RP2 (T-K test p=0.0032). The mean absolute Tat-induced activity of LTRs from RP2 was also found to be statistically significantly higher than both LTNP1 (T-K test p<0.0001) and LTNP2 (T-K test p=0.0314), however, although higher than the mean absolute Tat-induced activity of LTNP3 the mean activity of RP2 was not significantly higher than that of LTNP3 (T-K test p=0.0573). The mean absolute Tat-induced activity of LTRs from RP1, while higher, was not statistically higher than the mean activities of either LTNP2 (T-K test p=0.261) or LTNP3 (T-K test p=0.396).

As seen with the basal activities of the thirty LTR clones, examination of the data has indicated a non-significant (p=0.0595) trend in the Tat-induced activities of the same thirty LTRs within the Jurkat cell-line. The Tat-induced activities of the LTRs cloned from the RP patients in the analyses tended to be significantly higher within the Jurkat cell-line than the Tat-induced activities of the LTRs cloned from the LTNP patients.

Similar findings were also observed when the Tat responsiveness of each LTR clone was examined. The Tat responsiveness of each LTR clone corresponds to the fold-increase in LTR activity above basal levels that is directed by the HIV-2 LTR when in the presence of the HIV-2 Tat protein. Significant variation was observed between the Tat responsiveness of each of the thirty LTR clones (Figure 6.8, pg 242). In general, LTR clones that had directed the highest absolute levels of Tat-induced activity also showed the greatest fold increase in activity above basal levels. For example, the highest absolute level of Tat-induced activity was directed by the LTR clone RP3.3 (Figure 6.6A, pg 239). This clone also showed the greatest fold-increase in activity above its basal activity (Figure 6.8, pg 242), and was therefore the most active and the most responsive of the thirty LTRs in the presence of the Tat2 protein in this cell-line. However, this was not seen with every LTR clone. In particular, LTR clones from RP2 on average, directed higher absolute levels of Tat-induced activity than LTR clones from RP1 (Figure 6.6A, pg 239). When the mean Tat responsiveness of the LTRs from the two patients were compared, however, the LTRs from RP1 directed a higher fold increase in activity above basal levels on average (Figure 6.8, pg 242), and was therefore more Tat-responsive than the LTRs from RP2, even though the absolute activity of these LTRs in the presence of Tat2 was lower.

LTRs cloned from the three LTNP patients promoted a response to Tat that was on average 19-fold higher than basal activities, while the response directed by LTRs from the three RP patients was on average, 48-fold above basal activities. Even if the values corresponding to LTNP1, whose LTRs had not responded to Tat2, were removed from the analyses, the average Tat response directed by LTRs from LTNP2 and LTNP3 was 29-fold above basal levels, still considerably lower than the 48-fold observed on average with the LTRs derived from the three RP patients.

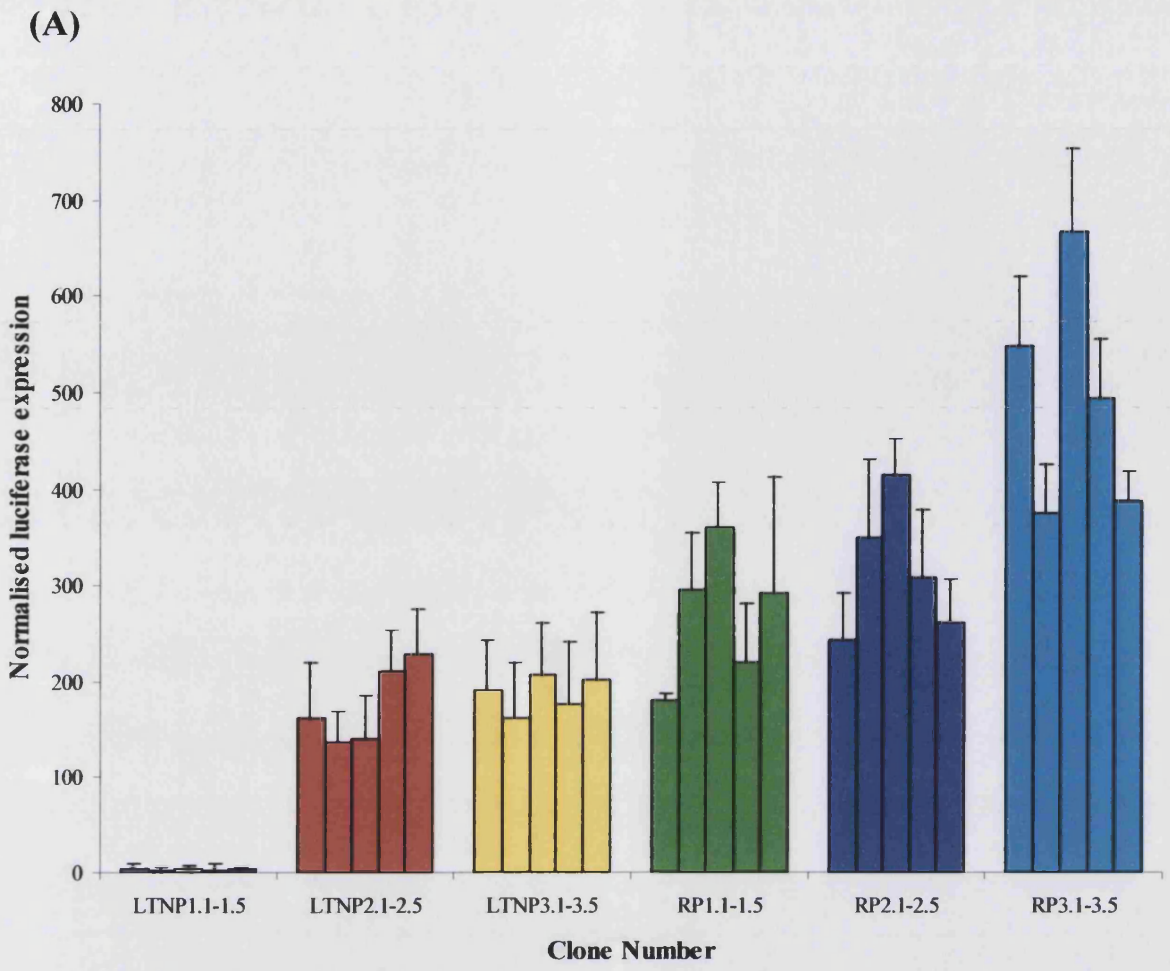


Figure 6.6 (A) Normalised absolute Tat-induced activities of 30 cloned LTRs in Jurkat cells (n=6). Five LTR clones represent each of six Gambian HIV-2 seropositive patients, three characterised as LTNPs, three as RPs. Each bar represents the mean normalised Tat-induced activity value calculated for each LTR clone from nine normalised Tat-induced activity values obtained in three independent triplicate cotransfections, with error bars representing the standard deviation of the nine individual values. Cotransfections were performed in 2 x 10^6 Jurkat cells with 1.25µg HIV-2LTR-pGL3E, 0.125µg of pRL-TK, 0.6µg of RSV-Tat and 0.025µg of HSDNA.

(B)

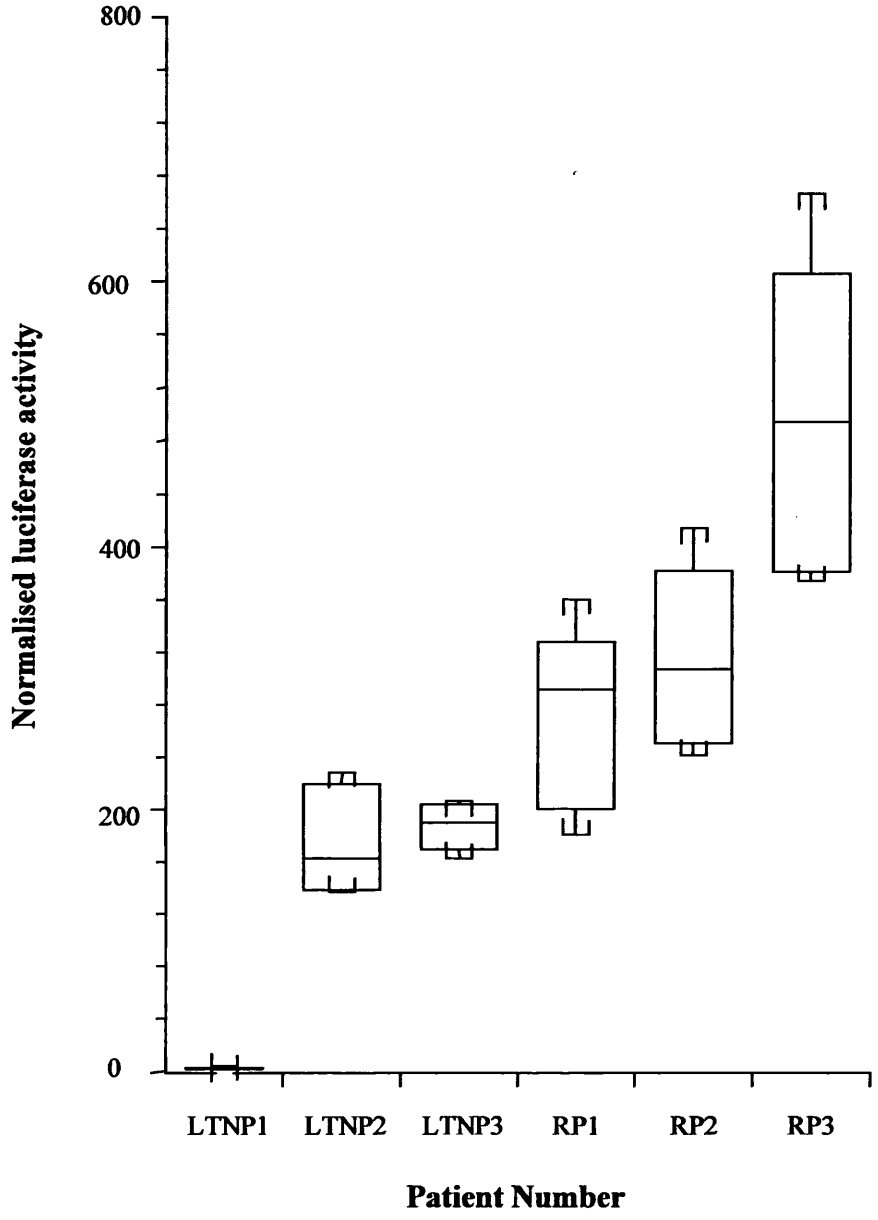
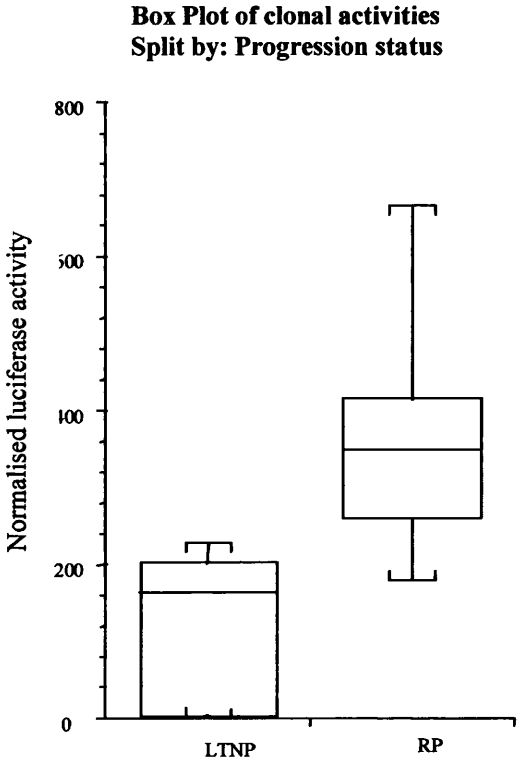


Figure 6.6(B) Box plot representing the range of absolute Tat-induced activities directed by the five LTR clones from each of six patients in Jurkat cells, the median Tat-induced LTR activity for each patient is marked by a line (n=6). Testing by one-way ANOVA revealed that the mean Tat-induced LTR activity of at least one of the six patients was statistically different from the other patients in the analyses ($p<0.0001$). Multiple comparisons by Tukey-Kramer tests revealed that the mean Tat-induced LTR activity of LTNP1 was significantly lower than all other patients both LTNP and RP ($p=0.0045$ – <0.0001). In addition, the mean Tat-induced LTR activity of RP3 was statistically significantly higher than all other patients both LTNP and RP ($p=0.0032$ - <0.0001). The mean Tat-induced LTR activity of RP2 was also found to be statistically significantly higher than LTNP1 ($p<0.0001$) and LTNP2 ($p=0.0314$).

(A)



(B)

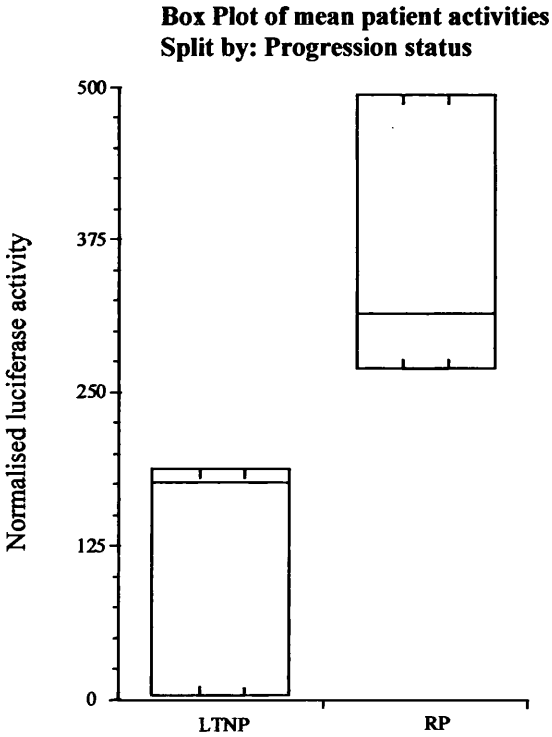


Figure 6.7 (A) Box plot representing the range of Tat-induced LTR activities directed by the fifteen clones from each progression group in Jurkat cells (n=6). Each progression group consisted of three patients from whom five LTRs were cloned. Characterisation of patients was based on clinical data available (Table 6.1, pg 223). Testing by one-way ANOVA revealed that the mean Tat-induced LTR activity of the RP group of clones was statistically higher than the mean basal LTR activity of the LTNP group of clones (n=30, $p<0.0001$). (B) Box plot representing the range of the three mean patient Tat-induced LTR activities in each progression group (n=6). Statistical comparison of the mean Tat-induced LTR activities of the two progression groups using the three patient means from each group (n=6) revealed that the difference between the mean Tat-induced LTR activities of the two groups became non-significant (p=0.0595).

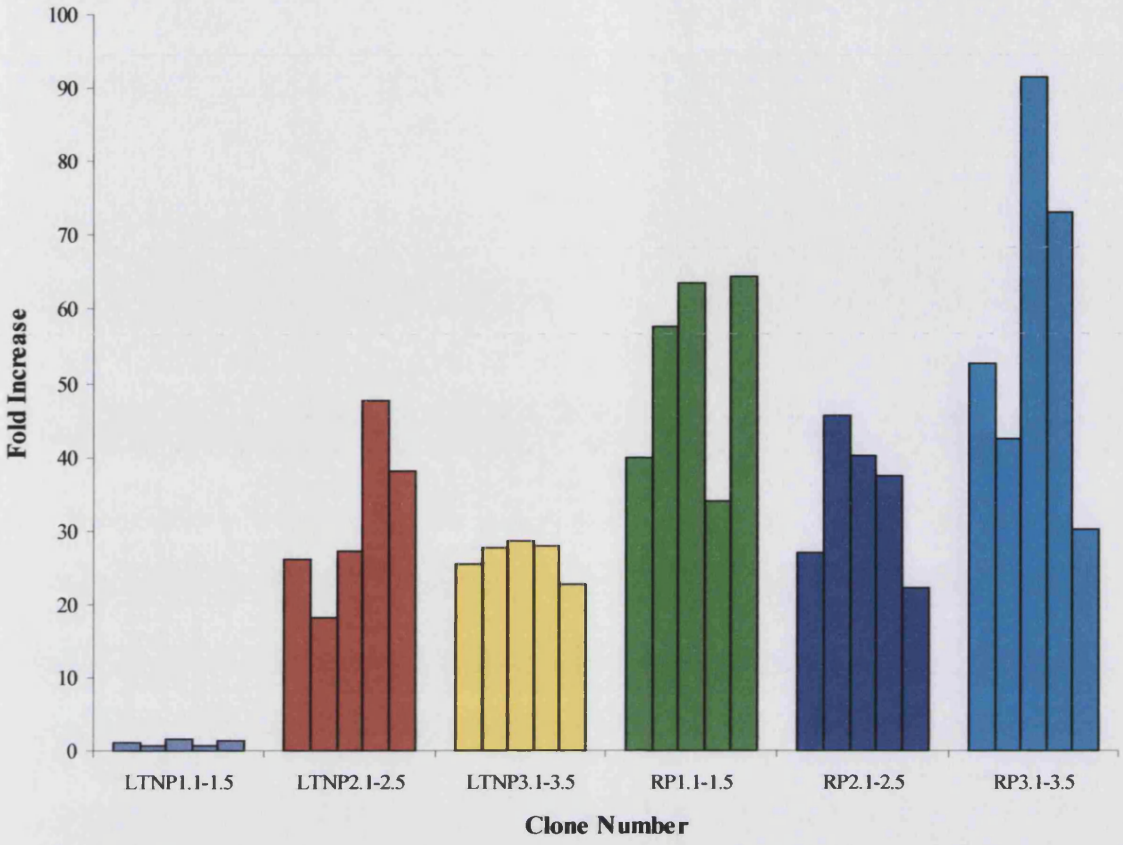


Figure 6.8 Mean Tat responsiveness of the LTRs cloned from each of the six Gambian patients (n=6). Tat responsiveness is measured as the fold increase in LTR activity following cotransfection with the pRSV-Tat expression vector and is calculated by dividing the mean normalised absolute Tat-induced activity of each LTR clone by the mean normalised basal activity of the same LTR clone.

6.3.3 Functional analysis of LTRs cloned from LTNP and RP patient PBMCs within the THP-1 cell line

LTR activity at the basal and Tat-induced levels was assessed in a second biologically relevant cell-line, the monocyte-like THP-1 cell-line. LTR function was assessed and compared using the same methodologies.

6.3.3.1 Basal activity of HIV-2 LTRs cloned from Gambian cohort patients in THP-1 cells

The basal activities directed by the thirty LTR clones, five from each of the six Gambian HIV-2 infected patients, LTNP1-3 and RP1-3, were assessed using the optimised dual luciferase assay system (see section 2.2.5). Each LTR clone was cotransfected in triplicate with the *Renilla* luciferase internal control vector pRL-TK at a ratio of 10:1 using the Effectene™ reagent protocol for the transfection of THP-1 cells (see section 2.2.4.6b). Over the course of the study this experiment was repeated a total of three times, and in parallel to the Tat activity analyses (see section 6.3.3.2). Luciferase activities derived from each of the HIV-2 promoters within the thirty LTR clones were normalised with respect to *Renilla* luciferase activity by dividing the firefly luciferase expression by the *Renilla* luciferase expression measured from each cotransfection (see section 2.2.5), and have been presented in the same format as used for the Jurkat data set (Section 6.3.2.1).

As seen at the basal level in Jurkat cells, when transfected into THP-1 cells each of the thirty LTR clones, taken from the three LTNP and three RP patients, directed a detectable level of basal activity (Figure 6.9A, pg 246). However, the overall levels of basal activity directed by each of the LTRs within the THP-1 cell-line were considerably less than observed within the Jurkat cell-line and activities fell over a narrower range.

Variation was observed between the basal activities of LTRs cloned from the same Gambian patient and between LTRs cloned from different Gambian patients. As observed in the Jurkat cell-line, the greatest intra-patient range of basal LTR activities was directed by LTRs cloned from RP3 (Figure 6.9A, pg 246, and B, pg 247). The range of activities directed by the LTRs from this patient was between two and four-fold wider than the range of activities directed by LTRs cloned from the other patients in the analyses.

A greater degree of overlap was observed between the basal activities of the thirty LTRs when transfected into THP-1 cells than had been observed between the activities of the same thirty clones when transfected into Jurkat cells. Overall, RP basal LTR activity ranged from 0.27 to 1.14, whilst the LTNP basal LTR

activity ranged from 0.13 to 0.91 (Figure 6.9A, pg 246, and Figure 6.10A, pg 248). The median basal LTR activities for the two progression groups were very similar at 0.48 (RP group) and 0.54 (LTNP group) (Figure 6.10A, pg 248). One-way ANOVA analysis of the basal activities of the thirty LTR clones when grouped as two progression groups (n=30), revealed that, unlike the basal activities of the same clones when transfected into Jurkat cells, the mean basal LTR activity of the LTNP group of clones was not significantly different (ANOVA p=0.57) from the mean basal LTR activity of the RP group of clones when transfected into THP-1 cells. Similarly, when this analysis was performed comparing the three LTNP patient mean basal activity values (calculated from the five clonal values for each patient) with the three RP patient mean basal activity values, no statistical difference was noted between the mean basal LTR activities of the two groups (ANOVA p=0.78) (Figure 6.10B, pg 248).

Of the thirty LTR clones tested, those cloned from RP3 directed the highest overall levels of basal activity within the THP-1 cell-line. LTRs cloned from this patient also directed amongst the highest basal activities when transfected into the Jurkat cell-line. The lowest LTR activities observed within the THP-1 cell-line were also directed by LTRs that had directed the lowest activities when transfected into Jurkat cells, and were cloned from LTNP1. However, cell-type specific transcriptional differences were noted between the basal activities of some LTR clones when transfected into the Jurkat and THP-1 cell-lines. In particular, LTRs cloned from RP2 had directed on average, the highest mean basal activity in Jurkat cells, with approximately three-fold higher activities than those directed by LTNP1. In the THP-1 cell-line LTR clones from RP2 directed on average the second lowest mean basal activity, the lowest belonging to the LTRs from LTNP1. The mean basal LTR activities of these two patients were indistinguishable within the THP-1 cell-line.

When the mean basal LTR activities of the six patients were individually tested and compared by one way ANOVA, the mean basal LTR activity of at least one of the six Gambian patients was found to be statistically significantly different to the mean basal LTR activities of the remaining patients in the analysis (ANOVA p=0.0016) (Figure 6.9B, pg 247). Multiple comparisons between the LTR activities of the individual patients by Tukey-Kramer tests revealed that the mean basal activity of LTRs from RP3 was significantly higher than the activities of LTRs from RP1 (T-K test p=0.013), RP2 (T-K test p=0.018), and LTNP1 (T-K test p=0.0023). The mean basal LTR activities of the remaining patients in the analyses were however, indistinguishable from each other statistically.

Thus in contrast to the findings in the Jurkat cell-line, examination of basal activities of the same thirty LTR clones in THP-1 cells has not revealed a trend in basal LTR activity. LTRs cloned from RP patients were not any more likely to be significantly more active at the basal level within the THP-1 cell-line than the LTRs cloned from LTNP patients in the analyses.

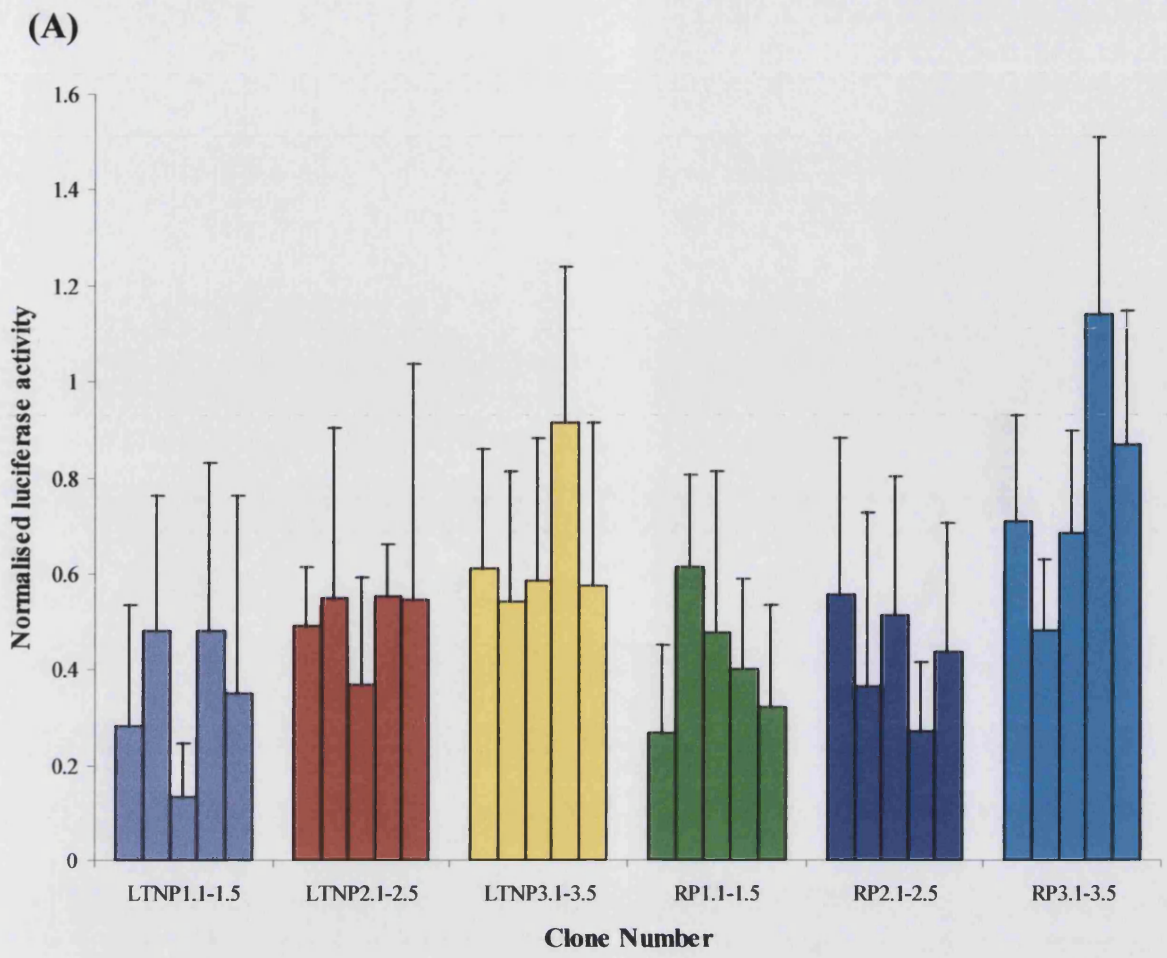


Figure 6.9 (A) Normalised basal activities of 30 cloned LTRs in THP-1 cells (n=6). Five LTR clones represent each of six Gambian HIV-2 seropositive patients, three characterised as LTNPs, three as RPs. Each bar represents the mean normalised basal activity value calculated for each LTR clone from nine normalised basal activity values obtained in three independent triplicate cotransfections, with error bars representing the standard deviation of the nine individual values. Cotransfections were performed in 2 x 10^6 THP-1 cells with 0.6μg HIV-2LTR-pGL3E, 0.06μg of pRL-TK and 0.14μg of HSDNA.

(B)

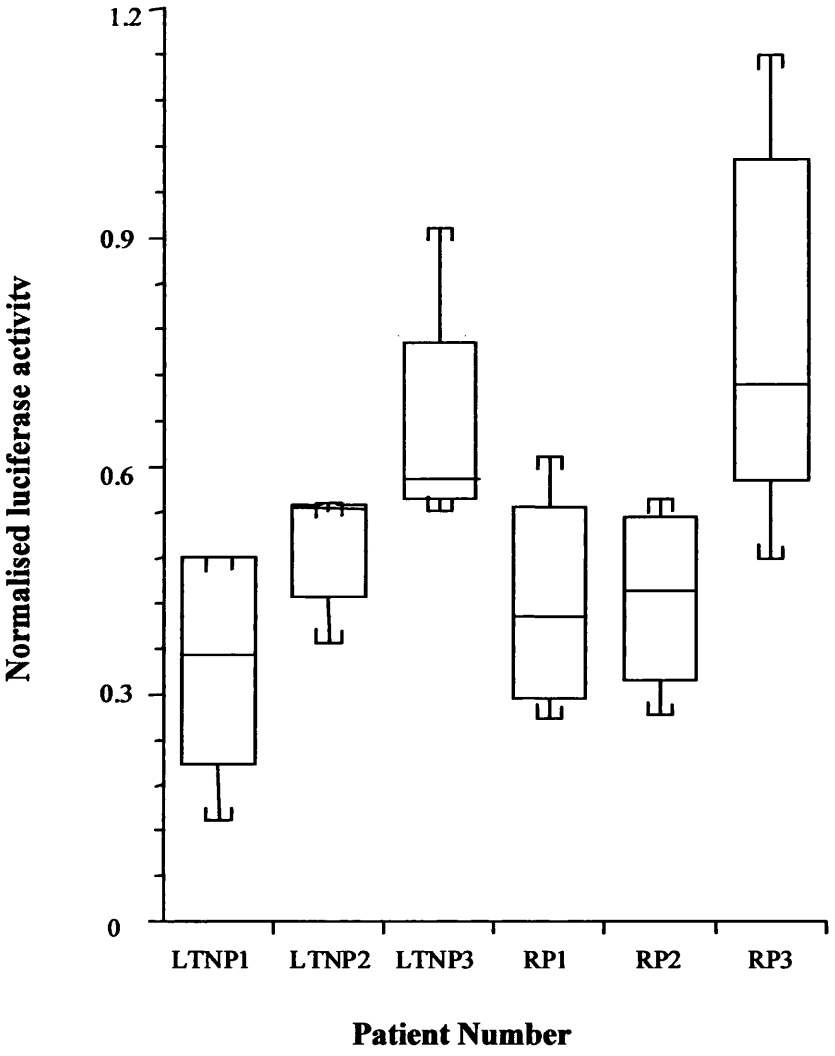
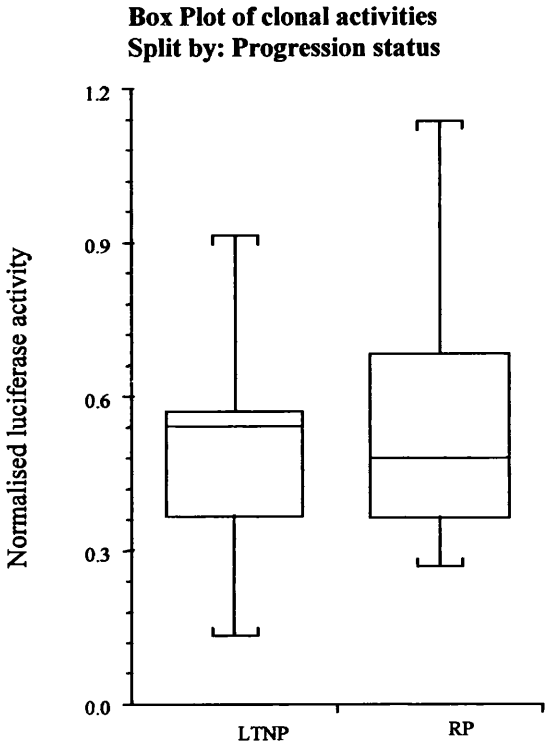


Figure 6.9(B) Box plot representing the range of basal activities directed by the five LTR clones from each of six patients in THP-1 cells, the median basal LTR activity for each patient is marked by a line (n=6). Testing by one-way ANOVA revealed that the mean basal LTR activity of at least one of the six patients was statistically different from the other patients in the analyses (p=0.0016). Multiple comparisons by Tukey-Kramer tests revealed that the mean basal LTR activity of RP3 was statistically significantly higher than the activities of LTRs from LTNP1 (p=0.0023), RP1 (p=0.013) and RP2 (p=0.0018).

(A)



(B)

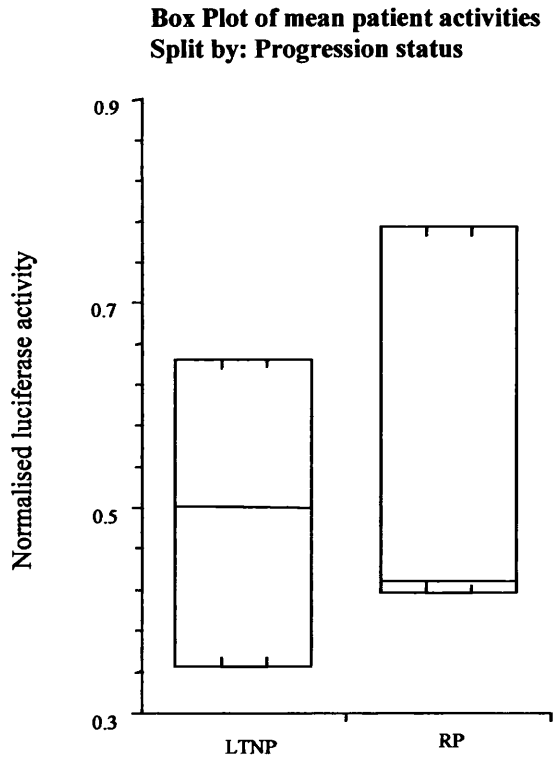


Figure 6.10 (A) Box plot representing the range of basal LTR activities directed by the fifteen clones from each progression group in THP-1 cells (n=6). Each progression group consisted of three patients from whom five LTRs were cloned. Characterisation of patients was based on clinical data available (Table 6.1, pg 223, and Figures 6.1, pg 224, and 6.2, pg 225). Testing by one-way ANOVA revealed that there was no statistical difference between the mean basal LTR activities of the RP and LTNP group of clones (n=30, $p<0.57$). (B) Box plot representing the range of the three mean patient basal LTR activities in each progression group (n=6). Statistical comparison of the mean basal LTR activities of the two progression groups using the three patient means from each group (n=6) also revealed no difference between the mean basal LTR activities of the two groups ($p=0.78$).

6.3.3.2 Tat-induced activity of HIV-2 LTRs cloned from the Gambian cohort patients in the THP-1 cell-line

As described in section 6.3.2.2, assessment and comparison of the transcriptional response to HIV-2 Tat directed by the HIV-2 LTR genotypes cloned from the three LTNP and three RP Gambian patients was achieved by including the HIV-2 Tat expression vector RSV-tat in the cotransfection mixtures for both Jurkat and THP-1 cell lines.

The thirty LTR clones, five from each of the six Gambian patients were cotransfected into the monocyte-like THP-1 cell-line with the *Renilla* luciferase internal control plasmid pRL-TK at a ratio of 10:1 and 125ng of the Tat2 expression vector pRSVtat. Cotransfections were performed in triplicate using the Effectene protocol described in section 2.2.4.7b. The experiment to determine the basal LTR activity of the Gambian LTR clones, (as discussed in section 6.3.3.1), was performed in parallel to the Tat experiment described below. Absolute Tat-induced LTR activities derived from each of the HIV-2 promoters within the thirty LTR clones were normalised with respect to *Renilla* luciferase activity by dividing the firefly luciferase expression by the *Renilla* luciferase expression measured from each cotransfection and have been presented in the same format as used for the Jurkat data set (Section 6.3.2.2). Over the course of the study, this experiment was repeated three times.

In addition, the normalised basal (Figure 6.9A, pg 246) and absolute Tat-induced activity (Figure 6.11A, pg 253) values have been used to calculate the Tat responsiveness of each LTR clone by dividing the normalised absolute Tat-induced activity by the normalised basal activity for each LTR clone. Results are presented as the mean fold increase in LTR activity above basal levels for each LTR clone in response to Tat2 (Figure 6.13, pg 256).

As seen in the Jurkat cell-line, with the exception of clones from LTNP1, the activity of each of the LTRs derived from the remaining five patients was increased significantly above basal levels when cotransfected with the Tat2 expression vector in THP-1 cells (Figure 6.11A, pg 253). The five LTR clones from LTNP1 that had appeared unresponsive to Tat2 in Jurkat cells also failed to direct any significant increase in activity when cotransfected into THP-1 cells with the Tat2 expression vector.

As noted in all previously described experiments considerable intra- and inter-patient variation was observed between the activities of the thirty LTR clones (Figure 6.11A, pg 253). Similar to findings at the basal level of activity within the THP-1 cell-line, the range of Tat-induced activities directed by the thirty LTR clones was much lower and narrower than the range of activities observed in the Jurkat cell-

line. Such differences in magnitude and range of LTR activities between different cell types have been well documented in previous HIV-1 LTR studies and are thought to result from differences in the concentrations of particular transcription factors within each cell type.

Despite the smaller overall range of activities, the highest levels of intra-patient variation were again observed between the activities of the LTRs cloned from the RP group of patients. As observed in Jurkat cells, the ranges of Tat-induced activity directed by the five LTRs from each of the three RP patients were higher than the ranges of Tat-induced activities directed by LTRs from both LTNP1 and LTNP 2 (Figure 6.11A, pg 253, and B, pg 254). Although, in contrast to the Jurkat data set, the range of activities directed by LTRs from LTNP3 was as great as the ranges observed for RP2 and RP3. Within the THP-1 cell-line the greatest intra-patient range of Tat-induced activities was directed by LTRs from RP1.

Of the thirty LTR clones analysed, LTRs cloned from RP3 directed the highest absolute Tat-induced activities within the THP-1 cell-line, while LTRs cloned from LTNP1 directed the lowest Tat-induced activities (Figure 6.11A, pg 253). This had also been the case when the thirty LTR clones were transfected into the Jurkat cell-line (Figure 6.6A, pg 239). Overall, LTRs cloned from the RP group of patients directed absolute levels of Tat-induced activity ranging from 5.80 to 19.84 with a median Tat-induced activity of 12.33 (Figure 6.11A, pg 253, and Figure 6.12A, pg 255). The median Tat-induced activity of LTRs from the LTNP group of patients was found to be significantly lower than that directed by LTRs from the RP group of patients at 6.40, with activities ranging from 0.29 to 15.68.

One-way ANOVA analysis of the absolute Tat-induced activities of the thirty LTR clones when grouped as two progression groups (n=30), revealed that within the THP-1 cell-line, as observed within the Jurkat cell-line (ANOVA $p<0.0001$), the mean Tat-induced LTR activity of the fifteen clones from three RP patients was statistically significantly higher (ANOVA $p<0.0022$) than the mean Tat-induced LTR activity of fifteen clones from three LTNP patients. On average RP LTR clones were twice as active at the Tat-induced level in THP-1 cells than the LTRs derived from the LTNP patients. However, when statistical analysis was performed comparing the three RP patient mean Tat-induced activity values (calculated from the five clonal values for each patient) with the three LTNP patient mean Tat-induced activity values, the difference between the mean Tat-induced activities of the two progression groups became non-significant (ANOVA $p=0.214$) (Figure 6.12B, pg 255).

Similar to findings at the basal level of LTR activity within both Jurkat and THP-1 cell-lines, cell-type specific transcriptional differences were noted between the Tat-induced activities of the LTRs when

transfected into the two cell-lines. For example, the LTR within clone LTNP3.1 was one of the least active promoters at the Tat-induced level when transfected into Jurkat cells (Figure 6.6A, pg 239) and yet directed one of the highest Tat-induced activities when transfected into THP-1 cells (Figure 6.11A, pg 253). In addition, LTR clones RP1.2 and RP1.3 had directed Tat-induced activities in Jurkat cells that were considerably higher than the activities of the LTR clones from LTNP3 (Figure 6.6A, pg 239). However, when transfected into THP-1 cells the Tat-induced activities of all five LTR clones from LTNP3 were higher than the activities of RP1.2 and RP1.3 (Figure 6.11A, pg 253).

Statistical analyses of the mean absolute Tat-induced LTR activities of the six Gambian patients when not grouped as two progression groups revealed that at least one of the patient means was statistically significantly different to the mean absolute Tat-induced LTR activities of the other patients in the analysis (ANOVA n=30, $p<0.0001$). Multiple comparisons between the mean Tat-induced LTR activities of each of the individual Gambian patients by Tukey-Kramer tests revealed that the mean absolute Tat-induced activity of LTRs from LTNP1 was significantly lower than the mean Tat-induced LTR activities of the remaining five Gambian patients (T-K test $p<0.0001$ – $p=0.0035$). In addition, as observed in the analyses of the Tat-induced activities within the Jurkat cell-line (section 6.3.2.2), the mean absolute Tat-induced activity of LTRs from RP3 was significantly higher than the mean Tat-induced LTR activities of all five Gambian patients, including RP1 and RP2, the two other patients within the RP group (T-K test $p<0.0001$ – $p=0.034$). While the mean absolute Tat-induced activity of RP2 was higher than the mean Tat-induced LTR activities of LTNP1 and LTNP2, the difference was only significant in the case of RP2 vs. LTNP1. Differences between the mean absolute Tat-induced activities of LTRs from the remaining patients in the analyses were not statistically significant.

As seen with the basal and Tat-induced activities of the thirty LTR clones within the Jurkat cell-line (Figure 6.4A, pg 232, and 6.6A, pg 239), examination of the THP-1 Tat-induced activity data has indicated a non-significant (ANOVA, $p=0.214$) trend in the Tat-induced activities of the same thirty LTRs within the THP-1 cell-line (Figure 6.11A, pg 253). The Tat-induced activities of the LTRs cloned from the RP patients tended to be higher within the THP-1 cell-line than the Tat-induced activities of the LTRs cloned from the LTNP patients in the analyses.

Similar findings were also observed when the Tat responsiveness of each LTR clone in THP-1 cells was examined (Figure 6.13, pg 256). Significant variation was observed between the Tat responsiveness of each of the thirty LTR clones, however, in general, LTR clones that had directed the highest absolute

levels of Tat-induced activity also showed the greatest fold increase in activity above basal levels. On average, LTRs cloned from the three LTNP patients promoted a response to Tat that was 11-fold higher than basal activities, while the response directed by LTRs from the three RP patients was on average, 24-fold above basal activities. Even after the values corresponding to LTNP1, whose LTRs had not responded to Tat2, were removed from the analyses, the average Tat response directed by LTRs from LTNP2 and LTNP3 was 16-fold above basal levels, still lower than the 24-fold observed on average with the LTRs derived from the three RP patients.

(A)

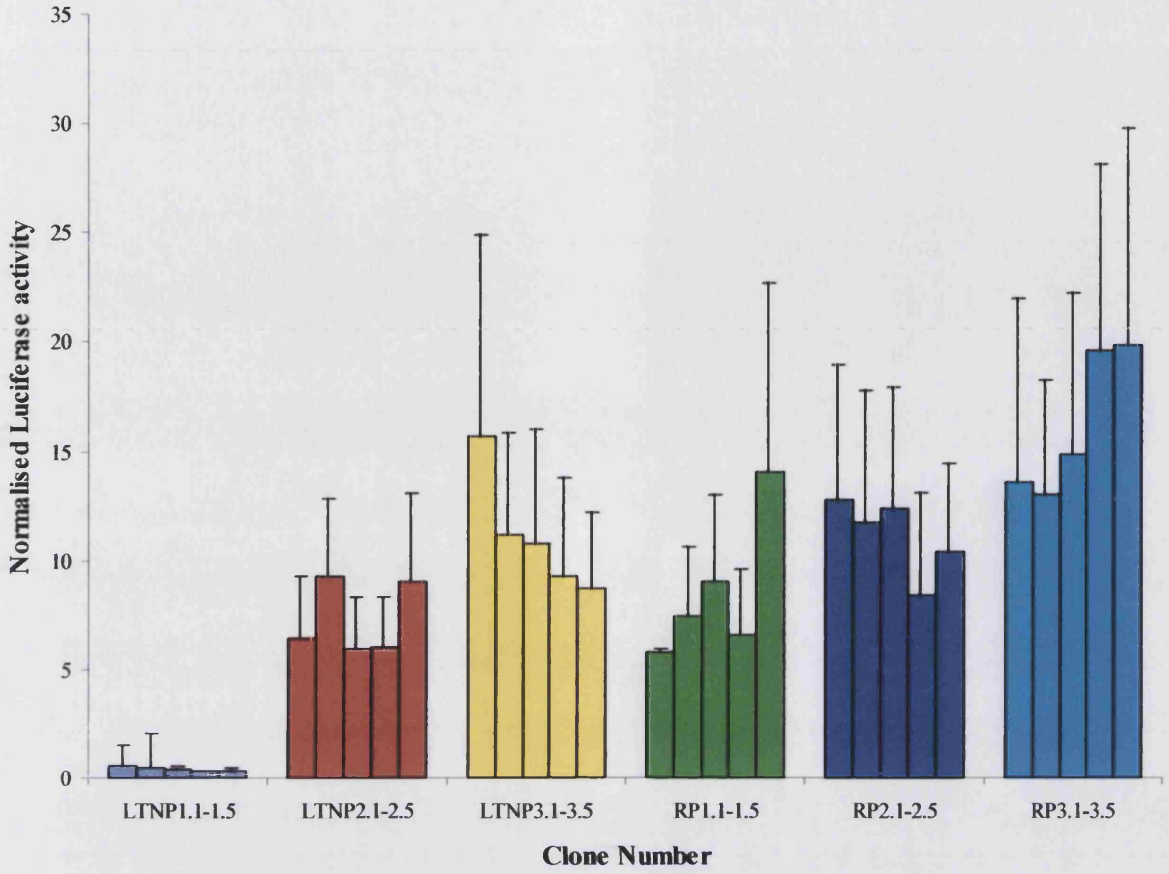


Figure 6.11(A) Normalised absolute Tat-induced activities of 30 cloned LTRs in THP-1 cells (n=6). Five LTR clones represent each of six Gambian HIV-2 seropositive patients, three characterised as LTNPs, three as RPs. Each bar represents the mean normalised Tat-induced activity value calculated for each LTR clone from nine normalised Tat-induced activity values obtained in three independent triplicate cotransfections, with error bars representing the standard deviation of the nine individual values. Cotransfections were performed in 2 x 10^6 THP-1 cells with 0.6µg HIV-2LTR-pGL3E, 0.06µg of pRL-TK, 0.125µg of RSV-Tat and 0.015µg of HSDNA.

(B)

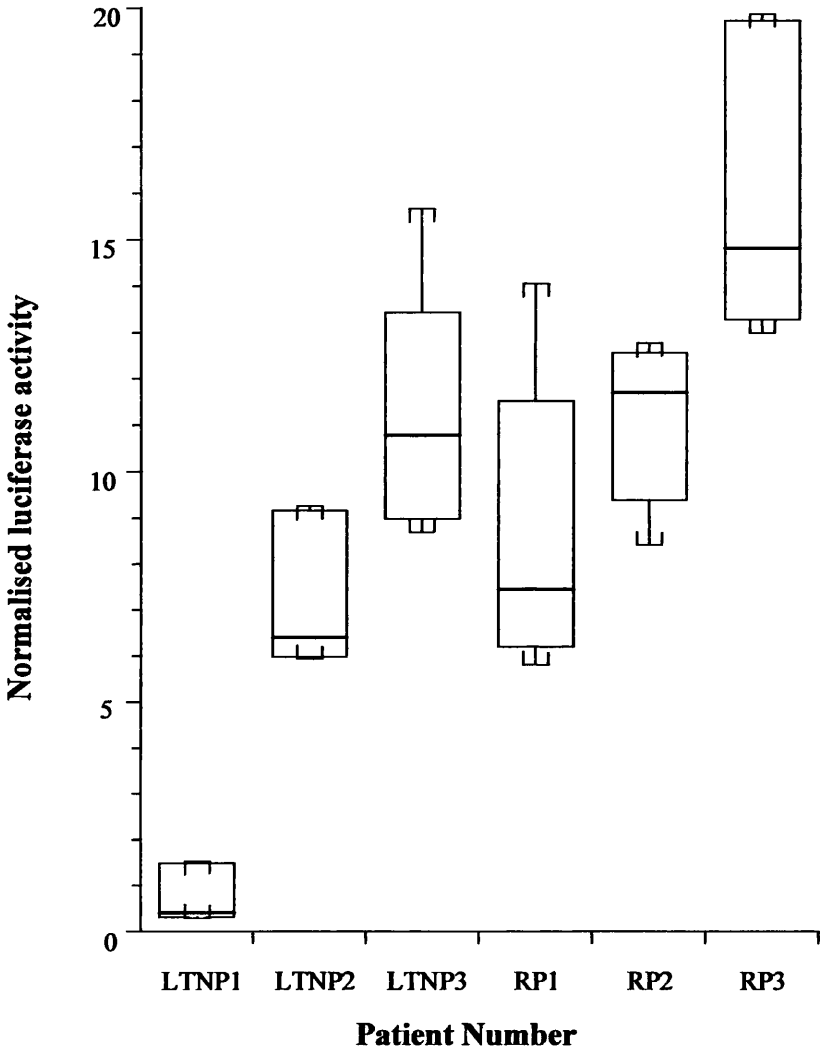
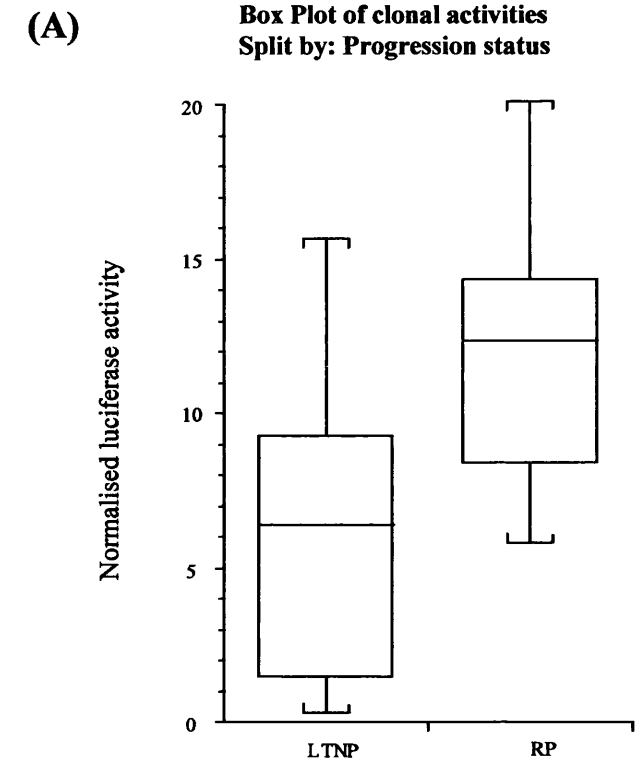


Figure 6.11(B) Box plot representing the range of absolute Tat-induced activities directed by the five LTR clones from each of six patients in THP-1 cells, the median Tat-induced LTR activity for each patient is marked by a line (n=6). Testing by one-way ANOVA revealed that the mean Tat-induced LTR activity of at least one of the six patients was statistically different from the other patients in the analyses ($p<0.0001$). Multiple comparisons by Tukey-Kramer tests revealed that the mean Tat-induced LTR activity of LTNP1 was significantly lower than all other patients both LTNP and RP ($p=0.0035$ – <0.0001). In addition, the mean Tat-induced LTR activity of RP3 was statistically significantly higher than all other patients both LTNP and RP ($p=0.0034$ - <0.0001).



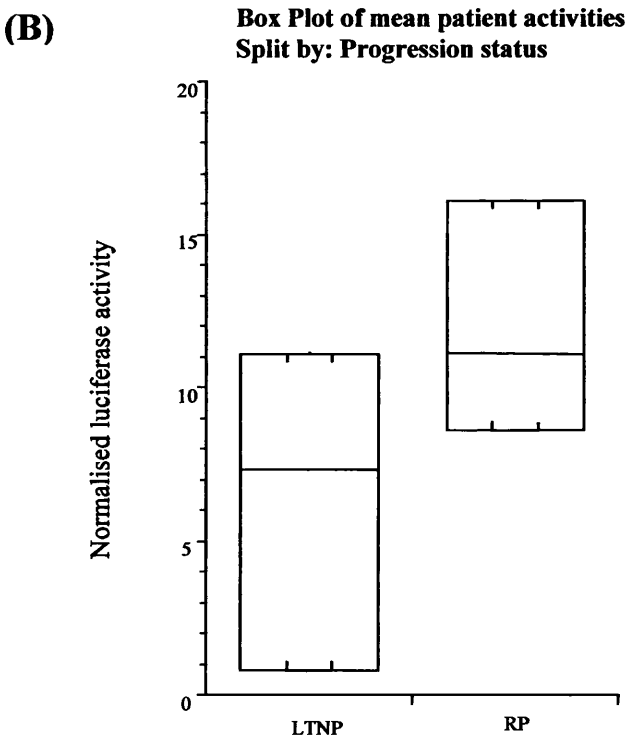


Figure 6.12 (A) Box plot representing the range of Tat-induced LTR activities directed by the fifteen clones from each progression group in THP-1 cells (n=6). Each progression group consisted of three patients from whom five LTRs were cloned. Characterisation of patients was based on clinical data available (Table 6.1, pg 223, and Figures 6.1, pg 224, and 6.2, pg 225). Testing by one-way ANOVA revealed that the mean Tat-induced LTR activity of the RP group of clones was statistically higher than the mean basal LTR activity of the LTNP group of clones (n=30, p=0.0022). (B) Box plot representing the range of the three mean patient Tat-induced LTR activities in each progression group (n=6). Statistical comparison of the mean Tat-induced LTR activities of the two progression groups using the three patient means from each group (n=6) revealed that the difference between the mean Tat-induced LTR activities of the two groups became non-significant (p=0.214).

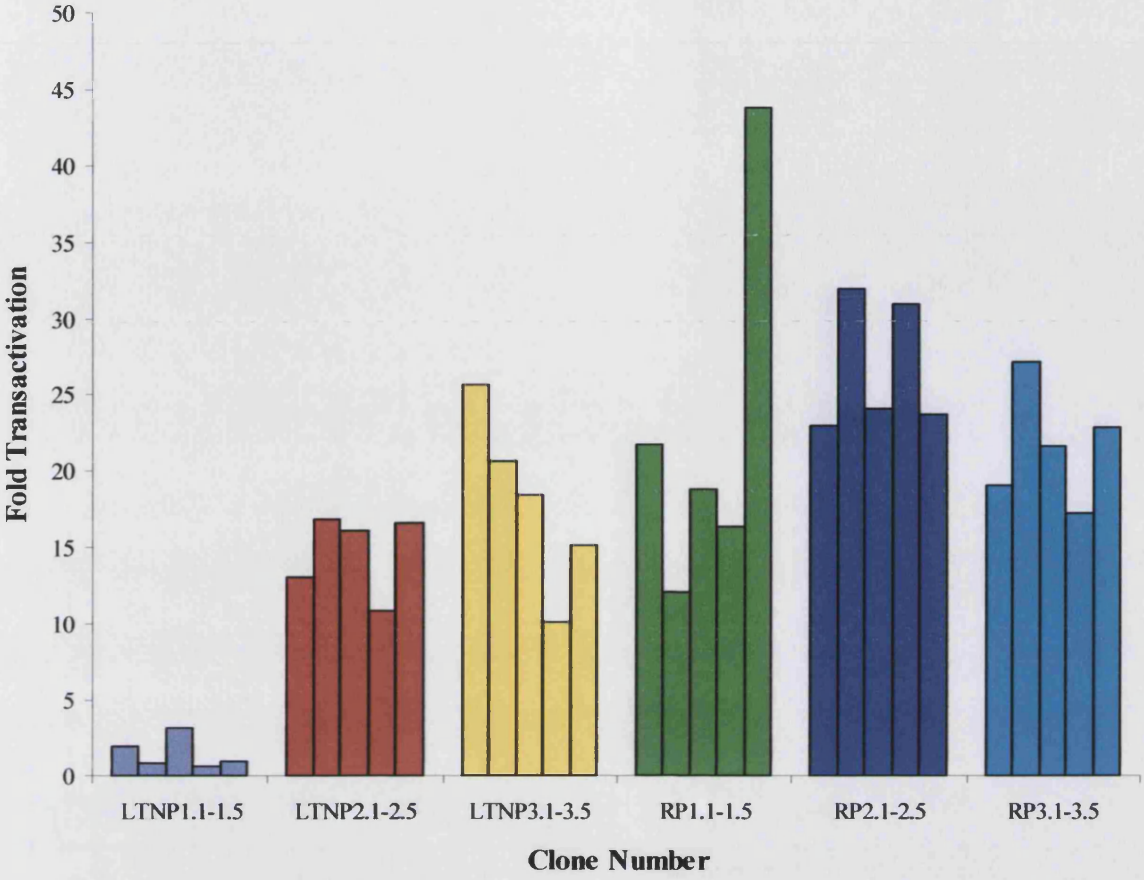


Figure 6.13 Mean Tat responsiveness of the LTRs cloned from each of the six Gambian patients (n=6). Tat responsiveness is measured as the fold increase in LTR activity following cotransfection with the pRSV-Tat expression vector and is calculated by dividing the mean normalised absolute Tat-induced activity of each LTR clone by the mean normalised basal activity of the same LTR clone.

6.4 Summary and discussion

In order to examine further the relationship between HIV-2 LTR activity and pathogenesis of HIV-2 infections, the basal and Tat-induced activities of LTRs cloned directly from the uncultured PBMCs of two contrasting groups of HIV-2 infected individuals were assessed and compared in two biologically relevant cell-lines.

The practicable number of LTR clones required to represent each of six Gambian patients in subsequent comparative transfection experiments was initially determined in an experiment performed to assess the intra-patient variation in basal LTR activity within the quasispecies of both a rapid progressor (RP2) and long-term non-progressor patient (LTNP3). Using the optimised dual luciferase assay system significant intra-patient variation was demonstrated between the basal activities of LTRs cloned from both the rapid progressor and long-term non-progressor viral populations. When transfected into Jurkat cells, the basal activities directed by the LTRs from RP2 fell over a 7-fold range whilst the basal activities of LTRs from LTNP3 fell over a slightly smaller range extending 6-fold. Statistically significant differences were noted between the activities of several of the LTR clones within both data sets. It was decided therefore, that multiple LTR clones would give a more representative view of naturally occurring promoter activity within the viral quasispecies of each Gambian HIV-2 infected patient. A practicable number of five LTR clones were chosen to represent each of the Gambian patients in the subsequent comparative transfection experiments.

6.4.1 Are LTNP LTRs completely defective?

When transfected into both Jurkat and THP-1 cells, the results demonstrated that each of the thirty LTR clones, five from each of the three rapid progressor and the three long-term non-progressor patients, were able to direct a detectable level of transcription at the basal level. Moreover, at the Tat-induced level of transcription the results from both cell-lines demonstrated that, with the exception of LTRs from LTNP1, the activities of the LTRs from five of the six Gambian patients increased significantly over basal levels, albeit to varying degrees, in the presence of the HIV-2 Tat protein. Since some degree of activity, either basal or Tat-induced, was detected from each of the thirty LTR clones, including the fifteen derived from the three LTNP patients, these data indicate that long-term non-progression in HIV-2 infection *per se* does not appear to be determined by the presence of completely inactive LTRs within the circulating quasispecies of these patients.

The low basal activity and lack of Tat responsiveness displayed by the LTRs from LTNP1 however, may reflect the existence of a viral population with significantly reduced promoter function circulating within this patient. Similarly reduced LTR function in a patient characterised as a long-term survivor (LTS) of HIV-1 infection has been reported by Zhang and co-workers (Zhang *et al.*, 1997b). Nucleotide sequence analyses of the LTRs from this patient revealed multiple G-to-A hypermutations within several of the core transcription factor binding sites, including the NF-κB, SP-1, and TAR elements. The functional effects of these mutations were severe and were proposed as the possible explanation for the long-term survival of this patient (Zhang *et al.*, 1997b). It was envisaged that nucleotide sequence analysis of the LTRs cloned from LTNP1 would also reveal significant polymorphism within the TAR and core transcription factor binding elements of these promoters, which could explain their severely reduced function. If this is the case, it may also be likely that the freedom from clinical illness experienced specifically by LTNP1 is related to the structural and functional defects of the promoter regions circulating within the quasispecies of this patient. Nucleotide sequence analyses of the LTRs cloned from LTNP1 were required to confirm this and are detailed in Chapter 7.

6.4.2 Variation in the function of naturally occurring HIV-2 LTR genotypes

In accordance with reports outlining the function of naturally occurring HIV-1 LTR genotypes (Michael *et al.*, 1994, Estable *et al.*, 1996, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b), a wide range of variation was observed between the activities of the thirty LTR clones analysed. Interestingly, significant differences were demonstrated between the intra-patient ranges of LTR activity from each patient in the analyses. Within THP-1 cells, and more particularly within Jurkat cells, the range of basal and Tat-induced LTR activities from each of the RP patients tended to be greater than the ranges of activities directed by the LTRs cloned from each of the three LTNP patients. In Jurkat cells, the ranges of basal LTR activity from both RP3 and RP2 were wider than the ranges of activities directed by LTR clones from LTNP1, 2 and 3. Moreover, the ranges of Tat-induced LTR activity from all three RP patients were, in each case, at least two-fold greater than the intra-patient ranges of Tat-induced activities directed by LTRs from each of the three LTNP patients in Jurkat cells and, both LTNP1 and LTNP2 in THP-1 cells. Notably, the range of basal and Tat-induced activities directed by the LTR clones from RP3 in Jurkat cells was not only greater than the ranges directed by LTRs from each of the three LTNP patients, but also greater than the ranges directed by the two other patients within the RP group.

Taken together, these data suggest that variation in promoter function exists within the quasispecies of each of the six Gambian patients in the analyses. Perhaps more importantly, they indicate that the level of variation in LTR function may differ between individuals, with a trend towards greater intra-patient variation in LTR activity within the quasispecies of RP patients being noted in the Jurkat cell-line, and, to a lesser extent within the THP-1 cell-line. Similar findings were noted in the analyses of the CBL isolates (Chapter 5), with the LTR activities of the more rapidly replicating isolates (CBL-20 and CBL21) that were derived from AIDS patients being more variable than the activities of the LTRs from the slower growing isolates that were derived from asymptomatic individuals (CBL-23 and CBL-24). Interestingly, this finding has not been reported by any of the HIV-1 studies that have compared the activities of LTRs from groups of LTS and RP patients (Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Quiñones-Mateu *et al.*, 1998). It is possible that the intra- and inter-patient variation observed in both the basal and Tat-induced LTR activities could result from transfection and/or detection system variation, however the likelihood of such variation being introduced is significantly reduced by the use of an internal control vector and the dual luciferase assay system. Moreover, there would be no reason for more variation to be introduced into the transfection and/or detection procedures used to assess the activities of the RP LTRs since both groups of LTR clones were treated in exactly the same manner within the same experiment. It is likely therefore; that the differences observed in intrapatient variation reflect genuine differences between the variation in LTR activity of the viral phenotypes circulating within each of the patients of the two progression groups, and indicate that promoter activity is more variable in patients exhibiting rapid progression to disease. It would be interesting to note whether the observed differences in the variation of LTNP and RP LTR function relate to differences in LTR sequence variation between the two progression groups. The relationship between LTR sequence and function is investigated in Chapter 7.

6.4.3 Comparison of LTNP and RP Basal LTR activities

Previous studies comparing the activity and nucleotide sequences of HIV-1 LTRs derived from long-term survivors (LTS) and typical progressors have failed to establish a general correlation between structural or functional LTR defects and disease progression (Estable *et al.*, 1996, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Quiñones-Mateu *et al.*, 1998, Visco-Comandini *et al.*, 1999, Gómez-Román *et al.*, 2000). A more recent study by Fang *et al* however, indicated that the transition from nonprogressive to progressive infection within an HIV-1-infected individual correlated with specific sequence changes in the LTRs circulating within that individual (Fang *et al.*, 2001). There appears to be the potential therefore, to select

for variation within the LTR that can lead to a more (or less) pathogenic phenotype. The data described below represents the first comparative functional analyses of the LTR region from LTNP and RP HIV-2-infected individuals.

Significant differences were noted between the overall basal LTR activities of the LTNP and RP groups when analysed, according to the cell-type transfected. Transfection experiments within Jurkat cells demonstrated that the RP LTRs tended to be more active at the basal level than the LTNP LTRs. Statistical comparison of the two progression groups showed that the mean basal activity of the fifteen RP LTR clones was significantly higher than the mean basal LTR activity of the fifteen LTNP LTR clones ($p<0.008$). However, when the analysis was reduced to comparing the mean basal LTR activities of each of the three LTNP patients with the mean basal LTR activity of each of the three RP patients, (calculated from the five clonal activities from each patient), the statistical difference between the two progression groups became non significant ($p=0.254$), making the overall trend observed in Jurkat cells statistically non-significant.

When analysed individually the mean basal LTR activities of two of the rapid progressor patients, RP2 and RP3, were found to be significantly higher in Jurkat cells than the mean basal activities of LTRs from two of the long-term non-progressor patients, LTNP1 ($p<0.0001$, $p<0.0001$) and LTNP2 ($p=0.0096$, $p=0.0138$), and higher though not statistically higher than the mean basal LTR activity of LTR clones from the third LTNP patient, LTNP3. Interestingly, the mean basal LTR activity of RP1 was found to be lower than the mean basal LTR activities of both LTNP3 and LTNP2, indicating that despite the general trend not all RP LTRs had higher basal activities than LTRs from LTNP patients in this cell-line. When tested statistically the activities of clones from LTNP1 were found to be significantly lower than the mean LTR activities of all patients in the analyses, with the exception of RP1.

While transfection experiments within THP-1 cells demonstrated that the LTRs cloned from RP3 directed the highest basal activities and the LTRs cloned from LTNP1 directed the lowest basal activities, overall the LTRs cloned from the rapid progressor patients in the analyses were not found to be any more likely to have a higher basal activity in THP-1 cells than the LTRs cloned from the long-term non-progressor patients. Statistical analyses performed with the THP-1 data set when grouped together as two progression groups revealed no statistical differences between the clonal ($n=30$) or the patient ($n=6$) mean basal LTR activities of the two progression groups in this cell-line ($p=0.67$ ($n=30$) and $p=0.78$ ($n=6$)).

Therefore, comparison of the basal activities of the thirty LTRs when transfected into THP-1 and Jurkat cells has revealed a difference between the patterns of activities directed by the LTRs in the two cell-lines. Not all of the LTRs that had displayed high levels of basal activity in one cell-line directed correspondingly high levels when transfected into the second cell-line. However, LTRs from LTNP1 and RP3 were found to direct the lowest and highest average basal activities respectively in both cell lines. The non-significant trend observed in the LTR activities of the thirty LTR clones when transfected into Jurkat cells provides additional evidence that some naturally occurring HIV-2 LTR genotypes can exhibit cell-type specificity in their function, and demonstrates more importantly that LTRs cloned from RP patients appear to exhibit transcriptional advantage at the basal level over LTRs cloned from LTNP patients in the T-cell-like Jurkat cell-line.

6.4.4 Comparison of LTNP and RP Tat-induced LTR activities

Examination of the Tat-induced activities of the thirty HIV-2 LTR clones revealed a similar but more pronounced trend in activity than that initially observed in the basal activities of these clones when transfected into Jurkat cells. On average, RP LTRs were found to be three times more active at the Tat-induced level in the Jurkat cells and twice as active in THP-1 cells than the LTNP LTRs. The activity of the RP LTRs in Jurkat cells was shown to increase on average 48-fold above basal levels in the presence of Tat compared to the average 19-fold increase observed with LTNP LTRs. Similarly, in THP-1 cells RP LTRs directed on average a 24-fold increase in activity above basal levels compared to the 11-fold average increase demonstrated by LTRs from the LTNP group of patients. Even following the removal of the data from LTNP1 the average LTNP Tat-induced LTR activity was still only found to be 29-fold above basal levels compared to the average 48-fold increase observed with RP LTRs in Jurkat cells, and 16-fold vs. the 11-fold average increase in THP-1 cells.

In both cell-lines, although most notably in Jurkat cells, the mean Tat-induced activities of the fifteen RP LTR clones was found to be statistically significantly higher than the mean Tat-induced activities of the fifteen LTNP LTR clones (Jurkat, $p<0.0001$, THP-1, $p<0.0022$). As observed at the basal level of activity however, when statistical analysis was performed comparing just the three patient mean Tat-induced activity values from each progression group, the difference between the mean Tat-induced activities of the two progression groups became non-significant in both cell-lines (Jurkat, $p=0.0595$, THP-1, $p=0.214$).

The basis for the observed general trend lies in the statistically significant differences noted between the mean Tat-induced LTR activities of the individual patients in the analysis. In both Jurkat and THP-1 cell-lines the mean Tat-induced LTR activities of two of the RP patients, RP2 and RP3, were found to be consistently and significantly higher than the mean Tat-induced LTR activities of two of the LTNP patients, LTNP1 and LTNP2. Where the trend was most pronounced, in Jurkat cells, the mean Tat-induced LTR activities of RP1, RP2 and RP3 were found to be higher than all three of the LTNP patients.

It is possible that the significant differences in the function of the LTRs from the RP and LTNP patients arise from differences between the core transcription factor binding elements of the LTRs from these different patients. In particular, throughout these functional analyses the LTRs from LTNP1 and RP3 have exhibited functional characteristics that distinguish them from the other patients in the analyses. In contrast to the severely reduced function displayed by the LTRs from LTNP1, the LTRs cloned from RP3 have directed activities at both basal and Tat-induced levels in both cell-lines that are either significantly and/or statistically higher than the activities of the LTRs cloned from the other LTNP and RP patients in the analyses. It is likely therefore, that the LTRs from RP3 have differences at the structural level as well as at the functional level from the other LTRs in the analyses. Nucleotide sequence analyses of the LTRs from this patient in addition to LTNP1 would therefore, be of interest and are detailed in Chapter 7. In addition, the LTRs from RP1 were consistently found to direct the lowest activities of the three RP patients in the analyses in both Jurkat and THP-1 cells. It is also possible that sequence differences exist between the core transcription factor binding elements of the LTRs from RP1 and those from RPs 2 and 3 that may account for the differences in function observed between these promoters.

In summary, we have shown that LTRs cloned directly from the PBMCs of patients exhibiting rapid progression to disease tended to have higher basal and Tat-induced activities when transfected into Jurkat cells than the LTRs cloned from patients exhibiting long-term non-progression. A similar but less pronounced trend was also observed in the absolute Tat-induced activities of the thirty LTRs when transfected into the monocyte-like cell-line, although overall the activities of the LTRs from the two contrasting groups of patients were more similar in THP-1 cells. The higher activities of the RP LTRs may reflect the existence of quasispecies within these patients with comparatively enhanced promoter activity. Conversely, the lower activities of the LTNP LTRs may reflect the existence of quasispecies within these patients with comparatively reduced promoter activity. By nature of the relationship between promoter activity, replicative capacity, and the rate of virus production, a higher rate of virus production and increased speed of symptomatic disease development could be predicted in an individual

containing a quasispecies of higher promoter activity when compared to an individual containing a quasispecies of low promoter activity. Significantly, the more pronounced difference between RP and LTNP LTR activity observed in Jurkat cells compared to THP-1 cells raises the possibility that the RP LTRs may provide transcriptional and replicative advantage within particular cell types.

The basis for the significant differences in basal and Tat-induced activity such as those observed between clones from RP3 and LTNP1 when transfected into Jurkat cells, and the more subtle overall differences observed between the mean basal and Tat-induced LTR activities of the progression groups as a whole when transfected into Jurkat and THP-1 cells, may lie in sequence differences between the transcription factor binding elements within the LTRs analysed. In characterising sequence changes to SP-1 binding sites in the HIV-1 LTR, McAllister *et al* demonstrated that substitutions within these sites resulted in greater transcriptional and replicative effects within the T-cell-like Jurkat cell-line when compared to the monocyte-like U937 cell-line (McAllister *et al.,* 2000). Similarly, Hilfinger *et al* demonstrated that changes to the *cis*-acting synergistic binding sites within the modulatory region of the HIV-2 LTR had a greater effect upon inducible transcription within immature monocytic cells such as HL-60s compared to the effects within mature monocytic cells such as THP-1s (Hilfinger *et al.,* 1993). In this vein, it is possible that the LTNP LTR sequences contain transcription factor binding site mutations that have a greater effect on LTR function in Jurkat cells when compared to THP-1 cells. In an attempt to determine a structural basis for the functional findings presented in this chapter, nucleotide sequence analysis was performed upon the thirty naturally occurring HIV-2 LTR genotypes. The findings of these analyses are presented in Chapter 7.

Chapter 7.

Further Characterisation of Naturally Occurring HIV-2 LTR Genotypes using Nucleotide Sequence Analyses.

7.1 Introduction

The analysis of DNA structure and its relationship to gene expression has been markedly facilitated by the development of powerful techniques for the sequencing of DNA molecules. Fluorescent dye-primer sequencing has been utilised in this chapter to determine the nucleotide sequence of each of the thirty cloned LTR genotypes described in the previous chapter. It is possible that the overall differences in activity observed between the LTNP and RP LTRs when analysed as two groups and, more probably, the comparatively low and high activities of LTRs from individual patients such as LTNP1 and RP3 respectively, could result from differences in the actual nucleotide sequences of the LTRs themselves. By determining and comparing the nucleotide sequences of the thirty cloned LTRs it was hoped that a greater understanding of the overall relationship between LTR structure, function, and disease progression status would be achieved.

7.2 Dye-primer cycle sequencing of naturally occurring HIV-2 LTR genotypes from LTNPs and RPs

7.2.1 Sequencing Oligonucleotides

Since the sequencing protocol described in section 2.2.6 routinely resulted in 400 accurately sequenced bases from the 3' end of the primer, to ensure the entire 764bp cloned LTR region was sequenced, three fluorescent dye labelled primers were designed. Two primers were designed to flank the cloned LTR region and were therefore located within the pGL3E firefly luciferase vector. The third primer was designed to bind to a sequence 317bp from the 5' end of the cloned LTR region; this conferred overlap of the flanking sequencing fragments and enabled a complete and accurate sequence to be obtained. The three overlapping but non-contiguous sequencing profiles were subsequently aligned into a contiguous sequence by downloading the computer-generated information into a sequencing programme called Sequencher (Gene Codes Corporation). All primers used in the sequencing reactions are listed in Table 2.6, pg 124.

7.3 Sequence analysis of LTR quasispecies from LTNP and RP patients

7.3.1 Rapid subtyping of the naturally occurring LTNP and RP LTR genotypes

Prototypic HIV-2 subtype A (HIV-2_{ROD}), B (HIV-2_{D205}) and simian-like strains (HIV-2_{F0784}) can be differentiated by a 40- to 44bp signature sequence (nucleotides –169 to –125) within the U3 section of the HIV-2 LTR (Berry *et al.*, 2001). The subtype of the LTR sequences from each of the six Gambian patients was determined by aligning a 30bp section of this signature sequence from each of the six patients with the corresponding sequence from prototypic HIV-2 subtype A, B and SIV-like strains (Figure 7.1, pg 266).

Each of the six Gambian LTR sequences, one sequence from each patient, were found to be phylogenetically most related to the HIV-2 strain belonging to subtype A, the most common HIV-2 subtype in West Africa (Figure 7.1, pg 266). Little homology was observed between the six Gambian LTR sequences and the viruses with subtype B (HIV-2_{D205}) or SIV-like (HIV-2_{F0794}) LTR sequences.

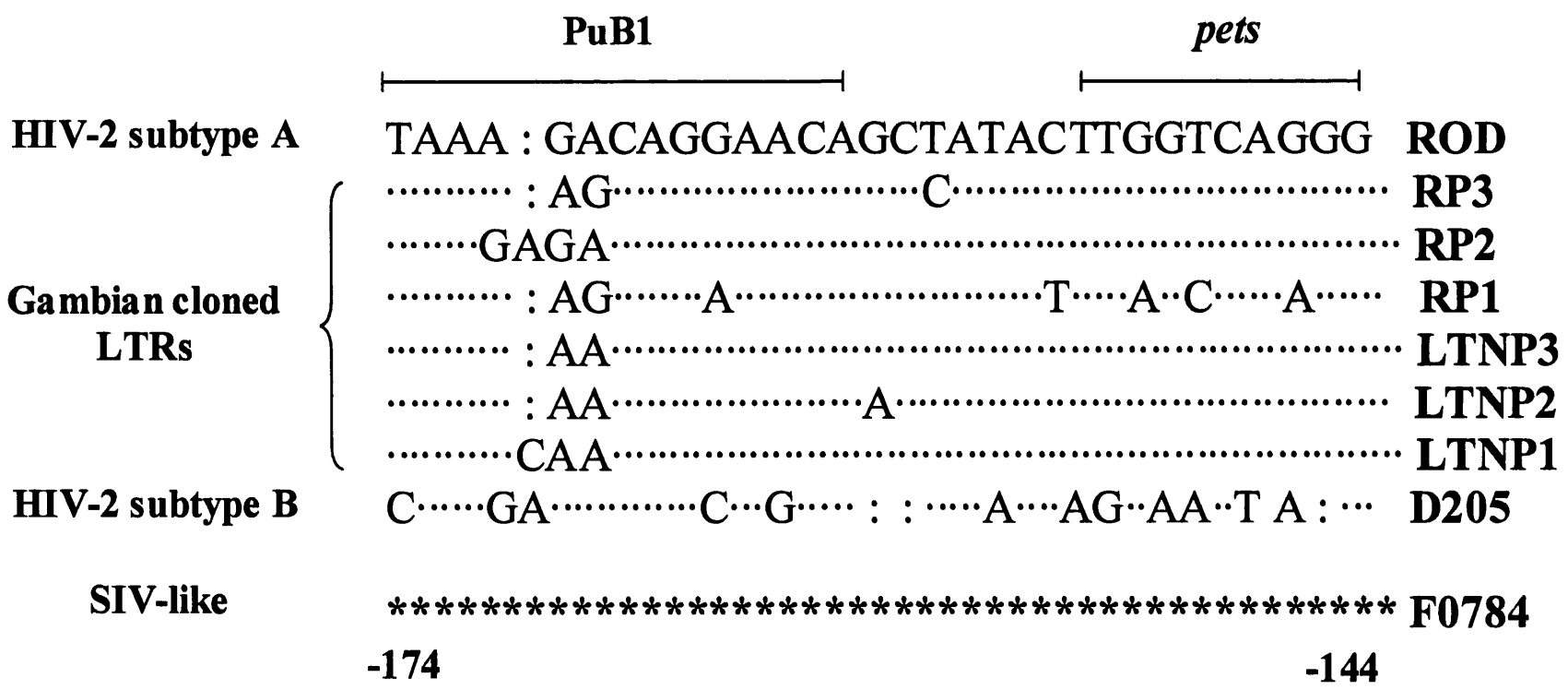


Figure 7.1 Alignment of six Gambian cloned LTR sequences (one from each patient) nucleotides –174 to -144, with a prototypic Subtype A (HIV-2_{ROD}), B (HIV-2_{D205}), and SIV-like (HIV-2_{F0784}) LTR sequence.

7.3.2 Nucleotide sequence variation in naturally occurring HIV-2 LTR genotypes from LTNP and RP patients

All thirty contiguous Gambian LTR sequences (five sequences from each of six patients) were then aligned and compared with reference to other HIV-2 subtype A sequences in the Los Alamos database (Los Alamos laboratory, 1997, Figure 7.2, pg 268) in order to determine genetic mutations. In addition, the degree of sequence diversity displayed by the cloned LTNP and RP LTR sequences was quantitatively analysed and compared using Kimura's two-parameter model (Section 7.3.3).

Comparison of the LTR sequences within the alignment (Figure 7.2, pg 268) revealed numerous nucleotide transitions throughout the cloned LTR sequences compared to the other prototypic subtype A HIV-2 sequences. All thirty contiguous sequence profiles were approximately the same length as the sequence profile of the prototypic subtype A HIV-2 strain ROD. This revealed that none of the LTR sequences cloned from either LTNP or RP patients contained any significant naturally occurring length polymorphisms, insertions, or deletions.

No characteristic sequence pattern distinguishing the long-term non-progressor LTR sequences from the rapid progressor LTR sequences was observed, however, when compared to the other reference sequences individual patient sequences could be distinguished from each other by characteristic patient specific sequence patterns within the alignment (Figure 7.2, pg 268)

Figure 7.2. Thirty LTR sequences, five from each of six Gambian HIV-2 infected individuals (LTNP1, 2, and 3)(RP1, 2 and 3) were aligned with reference to the consensus sequences of the LTR regions of subtype A HIV-2 and prototypic HIV-2_{ROD} from the HIV-2 sequence database from Los Alamos Laboratory. The sequences are ordered by patient identifier number at the beginning of each sequence, with RPs at the top and LTNPs at the bottom. The consensus of clade A and the HIV-2 $_{ROD}$ sequence is listed at the top of the alignment. A dash indicates identity with the sequence in consensus A. A semicolon indicates a gap introduced to preserve the alignment. Sites corresponding to transcription factor-binding sites (PuB1, PuB2, pets, Peri-κB, κB, SP1, TATA, TAR bulge and loop) are boxed. Bases highlighted in blue indicate a position where a novel mutation was observed.

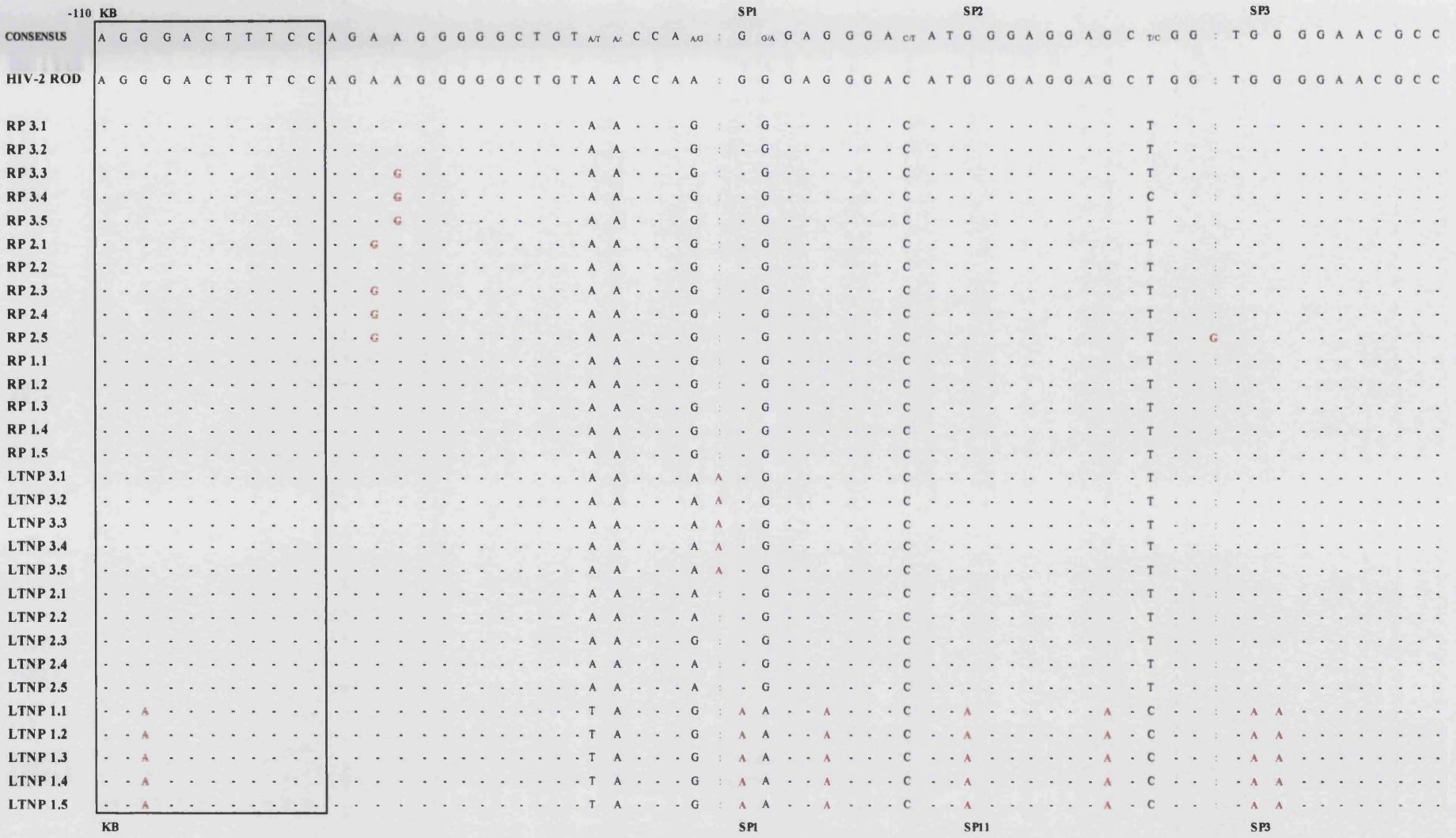


Figure 7.2. *Continued*

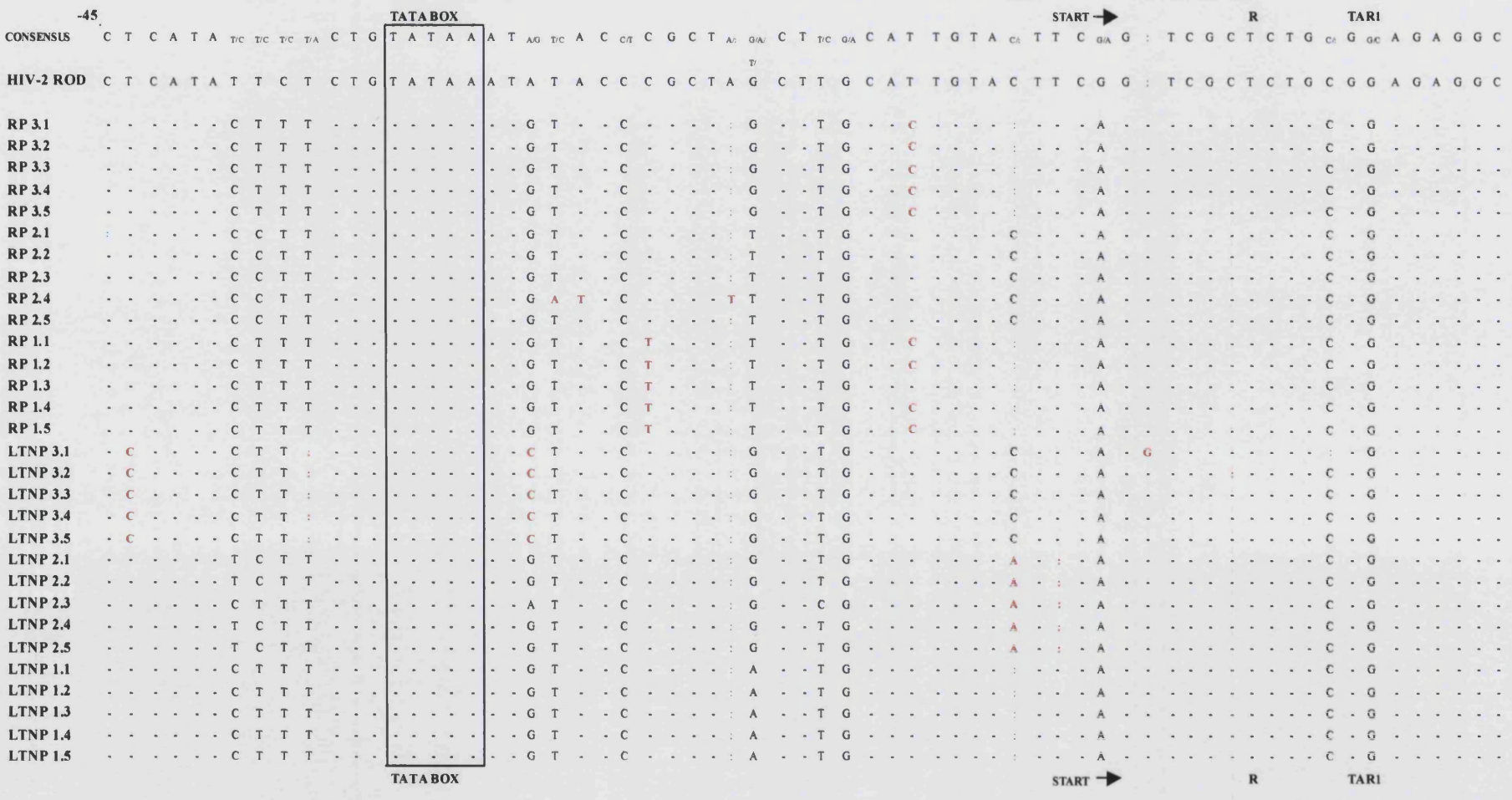


Figure 7.2. *Continued*

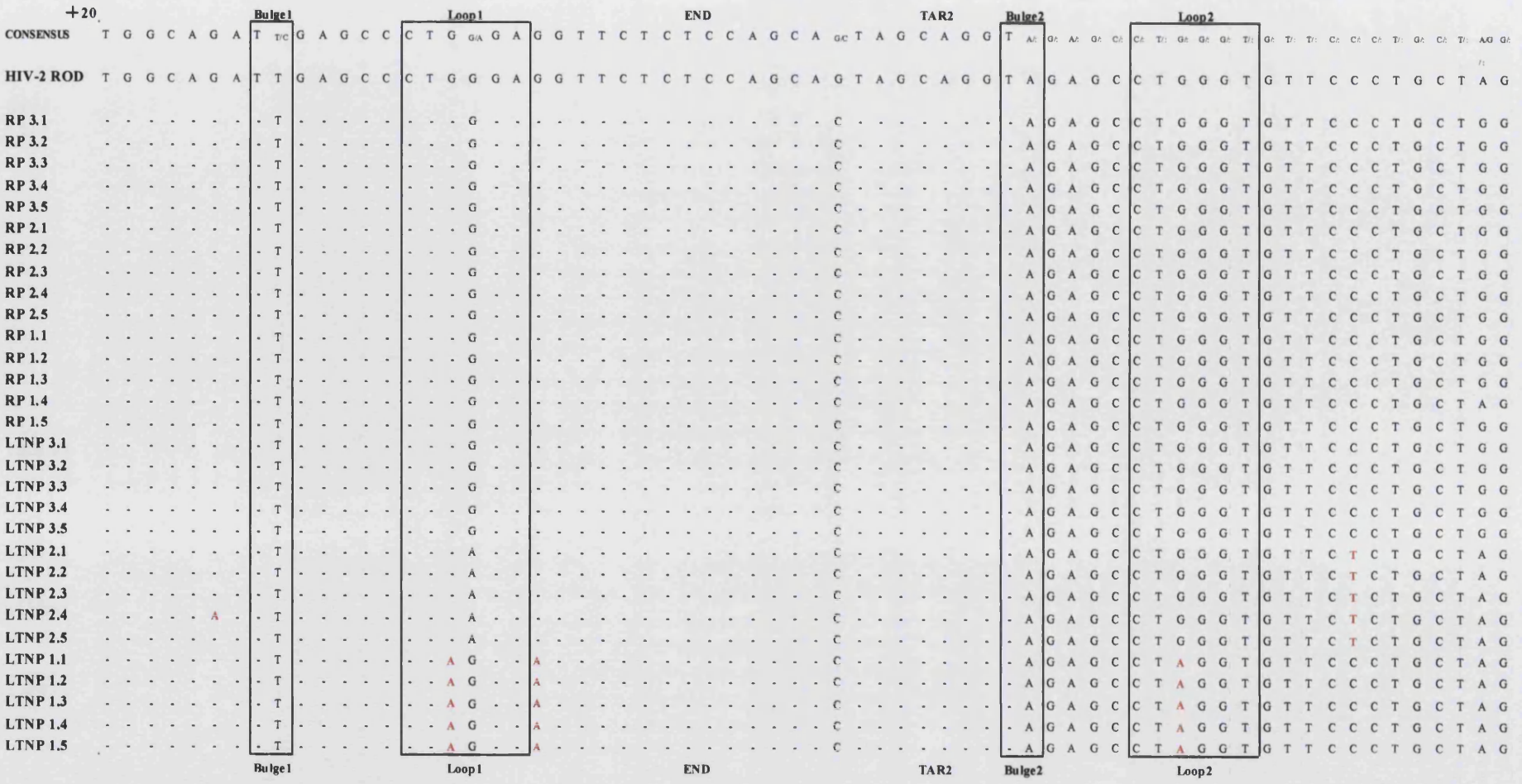


Figure 7.2. *Continued*

| | +86 | | |

7.3.2.1 In vivo HIV-2 LTR sequences from nucleotides -556 to -110

Analysis of the regulatory region of the HIV-2 LTR sequences revealed that some degree of sequence variation existed within the inducible *cis*-acting transcription factor binding sites located between nucleotides -175 to -110. When compared to the HIV-2 subtype A consensus, only 43% of the Gambian sequences had conserved PuB1 binding sites (nucleotides –174 to –161). Sequences from RP1, RP2, and LTNP1, bore single A or C base insertions at position 5 within the PuB1 site (Figure 7.2, pg 268). Notably sequences from RP1 also contained a G-to-A mutation at position 11 of the PuB1 site. The G-to-A change occurred within the 5'-AGGAA-3' pentanucleotide core of PuB1, the region of the site which acts as the specific binding motif for the Ets related proto-oncogenic transcription factor, Elf-1 (Leiden *et al.*, 1992, Markovitz *et al.*, 1992).

The second purine-rich site within the regulatory region of the HIV-2 LTR was more conserved among the Gambian HIV-2 LTR sequences. All sequences taken from the six study patients contained consensus PuB2 binding sites (nucleotides –142 to –136).

Eighty four percent of the Gambian LTR sequences had a strictly conserved 5'-TTGGTCAGGG-3' consensus sequence for the T-G rich pets site (nucleotides –152 to -144), the only variant found was 5'-TTAGCCAAGG-3', this variant was identified in 100% of sequences taken from RP1 (Figure 7.2, pg 268).

Several of the thirty Gambian LTR sequences displayed mutations within the 5' half of the peri-κB site (nucleotides –127 to -111), the inducible cell-type specific transcription factor-binding element of the regulatory region. LTR sequences from RP1 were found to contain a single G base insertion between the first two bases of the peri-κB site (Figure 7.2, pg 268). Sequences from LTNP3 showed a deletion of the G base located at position 1 of this binding site. The 3' half of the peri-κB site however, was found to be entirely conserved amongst the thirty LTR sequences.

Of the six patients analysed, only LTR sequences from RP3 and LTNP2 were found to have consensus sequences at all four *cis*-acting transcription factor binding elements within the regulatory region of the LTR. LTR sequences from RP1 were found to display the greatest number of nucleotide sequence changes within the regulatory *cis*-acting transcription factor binding elements, having mutations in the PuB1, pets, and peri-κB sites.

7.3.2.2 In vivo HIV-2 LTR sequences from nucleotides -110 to +1.

The NF-κB and SP1 binding sites located within the core region of the HIV-2 LTR remained largely conserved amongst the naturally occurring HIV-2 LTR sequences from both the LTNP and RP patients. A strikingly high number of G-to-A transitions were noted however, in both the NF-κB and SP1 binding sites of the sequences from LTNP1. All sequences derived from this patient displayed a G-to-A transition at position 2 of the NF-κB site (Figure 7.2, pg 268). In addition, multiple G-to-A hypermutations were observed within the central G-C rich core of all three SP-1 sites in 100% of all LTNP1 sequences.

The degree of sequence variation observed within the TATA box region of the Gambian LTR sequences was much less compared to that observed in the upstream *cis*-acting transcription factor binding elements in the regulatory region. The TATA box motif was conserved in all thirty LTR sequences when compared to the HIV-2 ROD sequence. However, all sequences derived from LTNP3 exhibited changes to highly conserved bases in the sequences flanking the TATA box. Sequences from this patient contained a T-to-C transition 5' to, and a G-to-C transversion directly 3' to the TATA box motif at positions -43 and –24 respectively (Figure 7.2, pg 268). In addition, all sequences from this patient contained a single base deletion at position –35, relative to the start site of transcription. The base changes at all three positions were not found in any HIV-2 or SIV sequences within the AIDS database.

HIV-2 initiator element sequence variants were identified in the LTR sequences from LTNP2. All sequences derived from this patient displayed a T base deletion at position -2 and an A base insertion at position -4 relative to the start site of transcription.

In contrast to the sequences from LTNP1, 2 and 3, LTR sequences from each of the three RP patients contained consensus sequences at each of the transcription factor binding elements within the core region of the HIV-2 LTR (nucleotides –110 - +1).

7.3.2.3 In vivo HIV-2 LTR sequences from nucleotides +1 to +198

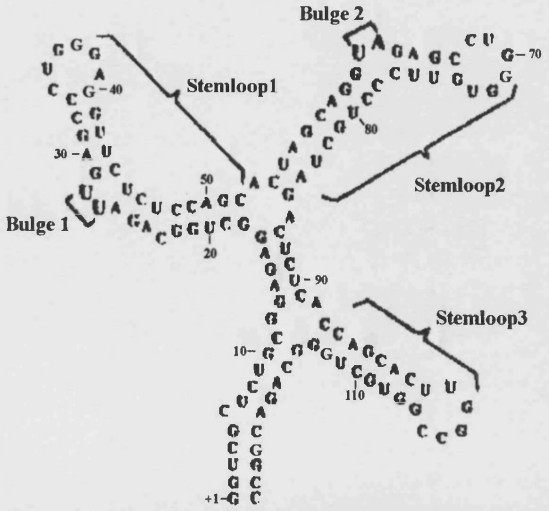
As with the core enhancer elements of the HIV-2 LTR, the degree of sequence variation within the TAR element of the thirty LTR sequences was much less in comparison to the variation observed within the upstream inducible transcription factor binding elements of the regulatory region. Changes to highly conserved bases within both stem-loop 1 and stem-loop2 of the TAR element were observed however, within the LTR sequences from LTNP1 and LTNP2 (Figure 7.2, pg 268, and 7.3, pg 276). The most

striking changes were observed within the LTR sequences from LTNP1. As observed in the NF-κB and SP-1 sites of this patient, multiple G-to-A hypermutations were also observed throughout the TAR region of all sequences from this patient. The most significant of those were located within the highly conserved hexanucelotide loop motifs of both stem-loop 1 and 2, the motifs of TAR that act as recognition sites for the cellular cofactor for Tat – cyclin T1 (Bieniasz *et al.,* 1998, Fujinaga *et al.,* 1998, Garber *et al.,* 1998a, Garber *et al.,* 1998b, Wei *et al.,* 1998, Richter *et al.,* 2002a, Richter *et al.,* 2002b). A G-to-A transition was observed at the third base within both loop motifs (+36 and +70 respectively from the transcriptional start site), changing the hexanucleotide loop sequence from 5'-CTGGGA-3' to 5'-CTAGGA-3' in stem-loop1, and 5'-CTGGGT-3' to 5'-CTAGGT-3' in stem-loop2 (Figure 7.3, pg 276). In addition, the G base at position +40 within the loop motif of stem-loop 1 had also been changed to an A base. The latter mutation is predicted to alter the base-pairing at the base of the unpaired loop in the first stem-loop structure, effectively increasing the number of unpaired residues within the loop from six to eight. Each of the five LTR sequences from LTNP-1 also contained G-to-A hypermutations at positions +112 and +120 of the TAR element. The mutation at +112 is located within the promoter-distal arm of stem-loop 3, a structure that has been shown to be functionally redundant in the Tat transactivation process (Fenrick *et al.,* 1989, Rhim & Rice, 1993). However, the G-to-A change at position +120 of TAR is predicted to affect the base pairing of the stem at the base of the TAR element, since the G base usually positioned at +120, basepairs with a C base located at position +4 of TAR (Figure 7.3, pg 276).

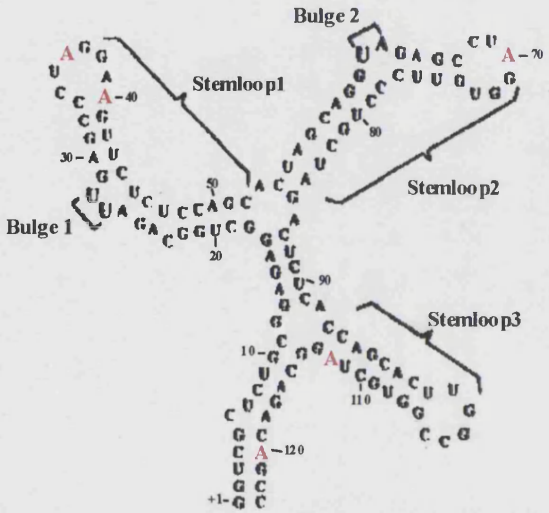
All sequences derived from LTNP2 showed a C-to-T base transition at position +78 of the TAR element (Figure 7.3, pg 276). This mutation occurs at a highly conserved base located on the promoter distal arm of stem-loop2. The C base located at +78 usually basepairs with a G base at position +61, directly below the Tat-binding bulge of stem-loop2. The mutation observed at position +78 in the sequences from LTNP2 is predicted to affect the base-pairing interactions within the Tat-binding bulge region of stem-loop2.

In contrast to the LTR sequences derived from LTNP patients 1 and 2, the TAR sequence elements of the LTR sequences derived from all three RP's and LTNP3 were entirely conserved and displayed no sequence variation from sequences within the AIDS database.

a) WT TAR



b) LTNP1 TAR



c) LTNP2 TAR

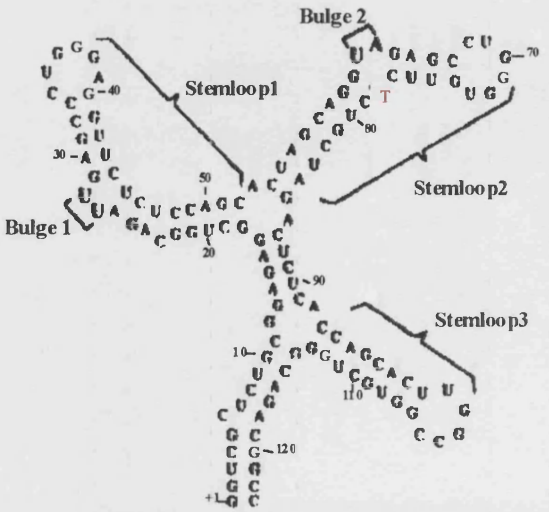


Figure 7.3 Schematic representations of the structures of a) wild-type HIV-2 ROD TAR RNA, b) LTNP1 HIV-2 TAR RNA and c) LTNP3 HIV-2 TAR RNA. Nucleotide bases highlighted in blue within both the TAR structures of LTNP1 and LTNP2 indicate a position where a novel mutation was observed.

7.3.3 Quantitative analyses of sequence diversity within the naturally occurring HIV-2 LTNP and RP LTR genotypes

To delineate further the genetic features of the Gambian HIV-2 LTR sequences, the degree of sequence diversity that the LTNP and RP LTRs showed from the consensus subtype A HIV-2 LTR sequence was quantitatively analysed and compared using Kimura's two-parameter model. Any deviation from the consensus A sequence was considered as a single mutation. In calculations of percentage variation, each duplication or deletion was counted as a single mutation. The G-to-A mutation rate was calculated as the number of G-to-A substitutions divided by the number of G sites. The general mutation rates and the specific G-to-A substitution rates displayed in the whole LTR sequence and at single LTR functional sites were determined for each of the thirty LTR sequences and are presented in Table 7.1, pg 279.

7.3.3.1 General mutation rates in the whole LTR sequence and at single LTR functional sites

When analysing the general mutation rate of the LTR as a whole, little difference was observed between the average mutation rates exhibited by the LTNP and RP LTR sequences (Table 7.1, pg 279). On average, LTNP LTRs showed 9.7% sequence diversity from the consensus subtype A sequence, while LTRs from the RP group of patients displayed 9.3% sequence diversity. Individual general mutation rates ranged from 8.8 to 11.0% for the LTNP group of patients and from 8.4 to 10.8% for the RP group of patients. Of the six patients analysed, LTR sequences from LTNP1 were found to display the highest rates of mutation (11.0%), while LTR sequences from RP3 exhibited the lowest rate of mutation (8.4%).

In contrast, analyses of the general mutation rate at single LTR functional sites within the LTNP and RP groups of LTR sequences revealed that the average rate of mutation at single LTR functional sites was significantly higher in the LTNP group of LTR sequences when compared to the average rate observed within the RP group of LTR sequences (2.8%, SD 0.95 (n=15) versus 1.1%, SD 1.82, (n=15); $p = 0.0024$) (Table 7.1, pg 279). On average, mutation rates at single LTR functional sites within LTNP LTR sequences were 2.5 times higher than the same rates observed within the RP LTR sequences. Individual patient average general mutation rates at single LTR functional sites ranged from 1.5% (LTNP3) to 5.3% (LTNP1) for patients in the LTNP group, and from 0% (RP3) to 2.11% (RP1) for patients in the RP group. Statistical comparison of the three patient averages from each progression group was found to reduce the statistical significance of the difference between the average general mutation rate at single LTR functional sites for the two progression groups from 0.0024 to 0.28. Of the six patients analysed, LTR sequences from LTNP1 were found to display the highest rate of mutation at single LTR functional

sites (5.3%), whilst LTR sequences from RP3 displayed consensus sequences at all single LTR functional sites, having an average general mutation rate of 0%.

7.3.3.2 G-to-A substitution rates in the whole LTR sequence and at single LTR functional sites

The average rate of G-to-A substitutions in the whole LTR was also found to be different for the two groups of LTR sequences (Table 7.1, pg 279). LTR sequences from the LTNP group of patients exhibited on average, twice the number of G-to-A substitutions observed in the RP LTR sequences. The average G-to-A substitution rate observed throughout the whole LTR for sequences from the LTNP group was 4.6% (SD 4.0), by comparison the same rate measured in the LTR sequences from the RP group of patients was on average, only 2.0% (SD 1.1) ($p = 0.022$ - n=30, $p = 0.41$ – n=6). Individual patient G-to-A mutation rates ranged from 1.2 to 10.1% in the LTNP group, with sequences from LTNP2 and LTNP1 displaying the lowest and highest rates respectively in this group, and from 0.7 to 2.9% in the RP group of patients, with sequences from RP3 and RP2 displaying the lowest and highest rates respectively in the RP group of patients.

Similarly, when G-to-A mutation rates at single LTR functional sites within both the LTNP and RP group of LTR sequences were calculated, LTR sequences from the LTNP group of patients were found to exhibit five times the number of G-to-A substitutions at single LTR functional sites compared to single LTR functional sites in RP LTR sequences (Table 7.1, pg 279). On average, the G-to-A substitution rate observed at single functional sites in the LTNP LTR sequences was 5.7%. By comparison, the mean G-to-A substitution rate observed at single functional sites in the RP LTR sequences was only 1.1%. The difference observed between the average mutation rates of the two groups of sequences however, was not found to be statistically significant ($p = 0.052$ – n=30). The individual G-to-A mutation rates at single functional LTR sites for the LTNP group of patients ranged from 0% (LTNP3) to 16.67% (LTNP1), and from 0% (RP2 and 3) to 3.3% (RP1) in the RP group. Only LTR sequences from RP1 of the RP group of patients were found to contain G-to-A substitutions within single LTR functional sites, the single LTR functional sites within the sequences from RP2 and RP3 failed to display any such substitutions. Strikingly, the G-to-A substitution rate observed at single functional sites in the LTRs from LTNP1 was found to be 16.7%, more than five times higher than the rate observed in any of the other LTR sequences analysed.

	General mutation rate in whole LTR (%)	*General mutation rate at single functional LTR sites (%)*	*G-to-A substitution rate in whole LTR (%)*	*G-to-A substitution rate at single functional LTR sites (%)*
LTNP1	11.0	5.3	10.1	16.7
LTNP2	8.8	1.6	1.2	0.3
LTNP3	9.4	1.5	2.6	0
LTNP Average	**9.7**	**2.8**	**4.6**	**5.7**
RP1	8.8	2.1	2.5	3.3
RP2	10.8	1.1	2.9	0
RP3	8.4	0	0.7	0
RP Average	**9.3**	**1.1**	**2.0**	**1.1**

Table 7.1 Average nucleotide variation in naturally occurring HIV-2 LTR genotypes cloned directly from LTNP and RP PBMCs. Analysis was performed using Kimura's two-parameter model. Any deviation from the consensus A sequence was considered as a single mutation. In calculations of percentage variation, each duplication or deletion was counted as a single mutation. The G-to-A mutation rate was calculated as the number of G-to-A substitutions divided by the number of G sites. The general mutation rates and the specific G-to-A substitution rates displayed in the whole LTR sequence and at single LTR functional sites were determined for each of the thirty LTR sequences and averaged for each of the six patients.

7.3.3.3 Intra and Interpatient sequence diversity

The extent of viral quasispecies heterogeneity within both the LTNP and RP group of patients was measured by calculating the intrapatient diversity exhibited by the LTR sequences from each patient in both progression groups, and the interpatient diversity displayed between sequences from each of the three patients in each progression group. Calculations were performed using Kimura's two-parameter model.

Analysis of the intrapatient diversity of the LTR sequences cloned from each of the six Gambian patients revealed a similar degree of diversity between the sequences from each of the patients in both groups. Individual intrapatient variation values for the three LTNP patients ranged from 0.3% (LTNP1 and 3) to 3.6% (LTNP2), with an average LTNP intrapatient variation value of 1.4%, and from 1.0% (RP3) to 3.8% (RP1) for each of the three RP patients, with an average intrapatient variation value of 2.0% for the RP group of patients. Of the six patients analysed the five LTR sequences from both LTNP1 and LTNP3 showed the least intrapatient nucleotide sequence variation (0.3%), whilst the five LTR sequences from RP1 displayed the greatest intrapatient nucleotide sequence variation (3.8%).

The pairwise comparison of DNA distances between the three patients within each of the two progression groups revealed that the average interpatient variation observed between the sequences from the three different patients in each progression group was not only similar but identical for the two progression groups. In both groups, the mean value of interpatient variation was 10.5%. Interpatient variation values for the three LTNP patients ranged from 7.5% (LTNP2 vs. LTNP3) to 12.0% (LTNP1 vs. LTNP3), and from 8.6% (RP3 vs. RP1) to 12.0% (RP2 vs. RP1) for the three RP patients.

7.3.3.4 In vivo HIV-2 LTRs consist of two uncoupled loci

Mutation rates in the *nef*-LTR overlap region (nucleotides -556 - -172) of the thirty LTR sequences were found to be significantly higher than in the untranslated region of the LTR sequences (nucleotides -171 - +198) (11.9 versus 6.9% on average). This nearly twofold-greater variability in the Nef-coding sequences compared with the LTR noncoding sequences has been extensively reported in previous HIV-1 LTR studies (McNearney *et al.*, 1995, Estable *et al.*, 1996), and here, as in those studies, suggests a stronger functional constraint for the latter region of the HIV-2 LTR, and suggests independently evolving loci.

7.4 Summary and discussion

Extensive analyses of the HIV-1 LTR in long-term survivors of HIV-1 infection have suggested that truncations and single point mutations in functional sites of the LTR are associated with impaired replication and attenuated pathogenesis (Fujii *et al.,* 1997, Rousseau *et al.,* 1997, Zhang *et al.,* 1997b, Fang *et al.,* 2001). In addition, a high G-to-A mutation rate in crucial sites of the HIV-1 LTR (Zhang *et al.,* 1997b) and in the *gag* gene (Huang *et al.,* 1998b) have been previously reported by others and suggested as a possible cause of clinical long-term non-progression. However, while structural and functional defects in the HIV-1 LTR have been linked to the well being of individual HIV-1-infected long-term survivors no correlation has been found between the HIV-1 LTR and long-term non-progression in general. In conjunction with detailed functional analyses, we have analysed and compared thirty Gambian LTR sequences from three LTNP and three RP patients in an attempt to examine further the relationship between the HIV-2 LTR and the pathogenesis of HIV-2 infection.

Alignment of the thirty Gambian LTR sequences with other prototypic subtype A HIV-2 LTR sequences revealed that all Gambian sequence profiles conformed to the HIV-2 subtype A consensus. None of the LTR sequences, from either LTNP or RP patients, contained any significant naturally occurring length polymorphisms, insertions, truncations, or deletions. Furthermore, we failed to identify any characteristic sequence pattern that distinguished a LTNP LTR sequence from a RP LTR sequence; sequence differences that were noted appeared to be patient-specific rather than progression group-specific. These findings are similar to those previously published in a number of HIV-1 LTR studies (Estable *et al.,* 1996, Rousseau *et al.,* 1997, Zhang *et al.,* 1997b, Quiñones-Mateu *et al.,* 1998, Visco-Comandini *et al.,* 1999, Gómez-Román *et al.,* 2000) and indicate that discrete differences in LTR sequence rather than significant LTR deletions or insertions are more likely to play a role in the determination of HIV-2 pathogenesis by the promoter region.

Analyses of the intra and interpatient sequence diversity for the six patients revealed that while the average interpatient variation was identical within each of the two progression groups (10.5%), on average the level of intrapatient variation was slightly higher within RP patients compared to LTNP patients (2.0% vs. 1.4% respectively). Initially, this finding appears to correlate with the results from the functional analyses, whereby the activities of LTRs from each of the RP patients were found to be more variable than the activities of LTRs from each of the LTNP patients. However, the RP patient that had directed the most variable LTR activities (RP3) was not found to display the greatest intrapatient

sequence variation, in fact, of the three RP patients RP3 exhibited the least intrapatient LTR sequence variation. It is likely therefore, that the different positions of the sequence changes between LTRs from the same patient rather than the actual number of changes between the sequences play a more important role in determining intrapatient variation in LTR function.

Consistent with other reports comparing the genetic diversity within HIV-1 LTS and TP LTR nucleotide sequences (Huang *et al.,* 1995a, Zhang *et al.,* 1997b, Quiñones-Mateu *et al.,* 1998, Visco-Comandini *et al.,* 1999), the general mutation rate for the entire HIV-2 LTR region was found to be similar for the LTNP and RP groups of patients (9.7% vs. 9.3% respectively). However, significant differences were observed between the two progression groups when the G-to-A mutation rate for the HIV-2 LTR as a whole and both the general and G-to-A mutation rates at individual functional sites within the HIV-2 LTR were compared. LTR sequences from the LTNP group of patients were found to exhibit on average two-fold higher (p = 0.022, n=30) rates of G-to-A mutations throughout the LTR sequence as a whole. Moreover, on average, LTNP LTRs displayed 2.5 times more general mutations at single functional sites in the LTR (p = 0.0024, n=30), and five-fold higher (p = 0.052, n=30) rates of G-to-A substitutions at single LTR functional sites, than observed in the RP LTR sequences. These results suggest a relationship between HIV-2 LTR sequence variation and disease progression status, and differ from the majority of studies where this parameter was found to be similar between the LTRs of LTS and TP of HIV-1 infection (Estable *et al.,* 1996, Zhang *et al.,* 1997b, Quiñones-Mateu *et al.,* 1998, Visco-Comandini *et al.,* 1999, Gómez-Román *et al.,* 2000).

Analysis of which single LTR functional sites were mutated amongst the sequences from the six patients revealed that in addition to the LTNP LTR sequences exhibiting higher average rates of mutation within the functional sites of the LTR, the sites that were mutated in these sequences were those located in the core and TAR regions of the LTR. LTR sequences from all three LTNP patients were found to contain point mutations within one or more of the core transcription factor binding elements involved in the recognition of general transcription factors required for transcription initiation. In addition, LTR sequences from two of the three LTNP patients displayed changes to highly conserved bases within the TAR element of the LTR. In contrast, the LTR sequences from the three RP patients contained 100% consensus sequences at all binding motifs within both the core and TAR regions of the LTR.

Of the core and TAR element mutations observed within the LTNP sequences, the most severe were located within the sequences from LTNP1. Multiple G-to-A hypermutations were observed within the

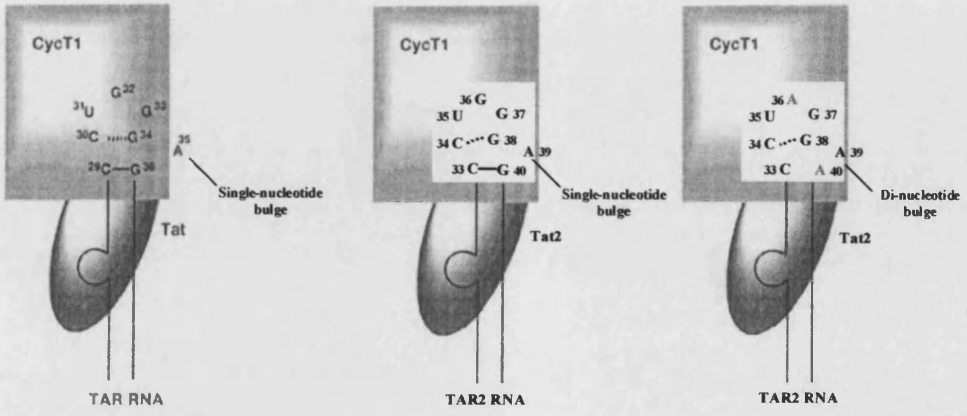
NF-κB, SP-1, and TAR binding sites of all LTRs cloned from this patient. The negative effects on viral transcription caused by mutations within conserved HIV-2 NF-κB (c Hannibal *et al.,* 1993, Hilfinger *et al.,* 1993, Hannibal *et al* 1994) and SP-1 sites (Arya & Gallo, 1988, Tong-Starksen *et al.,* 1990, Pagtakhan & Tong-Starksen, 1997) are well documented. It is probable that the multiple hypermutations observed within the NF-κB and SP-1 binding sites of the sequences from LTNP1 would severely reduce NF-κB and SP-1 protein binding, significantly impacting LTR function, and likely explain the low basal activity of the LTRs from this patient.

In addition, it is likely that the multiple G-to-A transitions exhibited within the TAR hexanucleotide loop motifs of stem-loop 1 and 2 explain the failure of the LTR sequences from LTNP1 to respond to Tat2 in any way in functional studies. In determining the specific HIV-1 TAR loop sequence and functional groups required for CycT1-Tat-TAR ternary complex formation Richter *et al* have demonstrated that the guanidine bases located at position 32 and 34 within the HIV-1 hexanucelotide loop are essential for CycT1-Tat binding to TAR RNA (Richter *et al.,* 2002a, Richter *et al.,* 2002b) (Figure 7.4, pg 284). The guanidine base at position 32 has an O6 hydrogen bond acceptor from a carbonyl group that has been shown to be critical for the interaction between the CycT1-Tat complex and TAR. By substituting G32 with an A base Richter *et al* demonstrated that CycT1-Tat binding was reduced to less than 40% capacity (Richter *et al.,* 2002a). The sequences from LTNP1 contain a G-to-A mutation at the G base located in the position equivalent to G32 in the HIV-2 loop motifs of both stem-loop 1 (G36) and stem-loop 2 (G70) (Figure 7.4, pg 284). It is likely that these G-to-A changes result in a similar reduction in CycT1-Tat2 binding to TAR. This in turn would result in a reduction in the level of transcription directed by the LTRs from LTNP1 in the presence of Tat2. Moreover, since both stem-loop1 and stem-loop2 contain this mutation, the second stem-loop structure would not be able to perform its well-documented partial rescue of TAR function in the context of a mutated stem-loop1 (Fenrick *et al.,* 1989, Berkhout *et al.,* 1990, Rhim & Rice, 1993, Garcia-Martinez *et al.,* 1995, Browning *et al.,* 1997).

The explanation for the complete abrogation of Tat-induced transcription directed by the LTRs from LTNP1 likely lies in the fact that the loop motif in stem-loop1 of LTNP1 also contains a G-to-A change at the base of the loop (G40), at the position equivalent to G36 in the HIV-1 TAR loop. The G at this position usually basepairs with the C base in the opposite arm of the stem and forms the first basepair below the hexanucelotide loop motif (see Figure 7.4, pg 284). The substituted A base at this position in stem-loop 1 of LTNP1 would effectively increase the loop motif from hexanucelotide to octanucleotide,

since the A would be unable to basepair with the C base. In characterising the interactions between the bases within the HIV-1 loop motif and the CycT1-Tat complex, Richter *et al* suggested the presence of a putative base pair between the first and fifth bases of the hexanucelotide loop, creating a single A base bulge at position 35 (Richter *et al.*, 2002a, Richter *et al.*, 2002b) (Figure 7.4, pg 284). It has been proposed that the presence of this bulge is important for distortion of the backbone RNA structure and the enhanced accessibility of the functional groups on G32 and G34 involved in the recognition of the CycT1-Tat protein complex (Richter *et al.*, 2002a, Richter *et al.*, 2002b). Since the sequences from LTNP1 contain a change to the G base at position 40 at the base of the loop, the single-nucleotide A base bulge would be changed to a dinucleotide A base bulge (Figure 7.4, pg 284). Based on the importance of the structure and sequence of this region of the motif to the CycT1-Tat-TAR interaction, it is likely that this change disrupts the accessibility of the CycT1-Tat2 complex to the functional groups involved in TAR2 binding and contributes significantly to the abrogation of Tat response directed by the LTRs from this patient.

Finally, HIV-2 Tat-SP1 transcriptional synergy has been reported (Blau *et al.*, 1996, Pagtakhan & Tong-Starksen, 1997). Since all three SP1 sites, in addition to the TAR element within the LTR sequences from LTNP1 are mutated, it is unlikely that such transcriptional synergy is possible in this case. Taken together, it is likely that the multiple G-to-A hypermutations throughout the LTR sequences from LTNP1 account for the dramatically reduced basal and completely abrogated Tat *trans*activated activity of the LTRs from this patient. It is likely that these findings reflect the presence of a viral population with a severely attenuated replicative capacity circulating within LTNP1 and suggests an explanation for the long-term non-progression experienced by this patient.



HIV-1 wt **HIV-2 wt** **HIV-2 LTNP1**

Figure 7.4 Schematic representation of the CycT1-Tat-TAR ternary complex for HIV-1 wt, HIV-2 wt, and HIV-2 LTNP1 (adapted from Richter *et al.*, 2002a).

All LTR sequences taken from LTNP2 were found to contain a single T base deletion and a single A base insertion at positions –2 and –4 respectively, relative to the start site of transcription. The nucleotides within this region of the LTR encode an initiator element that matches the loose consensus sequence for mammalian Inr activity of $YCA^{+1}NTYY$ where Y is a pyrimidine and N is any nucleotide (Javerhery *et al.,* 1994). Mutation of the base at position -2 within a mammalian Inr has been shown to reduce promoter strength by approximately two-fold (Javerhey *et al.,* 1994). Similarly changes within the core of the HIV-1 Inr element have been shown to dramatically reduce transcription initiation rates and limit Tat responses (Zenzie-Gregory *et al.,* 1993, Rittner *et al.,* 1995). Furthermore, changes to nucleotides between bases –6 to +30, encompassing the Inr consensus sequence of HIV-2 have been shown to reduce the transcriptional activity of the HIV-2 promoter (Jones *et al.,* 1988).

LTR sequences from LTNP2 were also found to display a C to T transition at position +78 within the second stem-loop region of the TAR element. Since residue +78 basepairs with the G residue at position +61 at the base of the second trinucleotide bulge, the change to a T base at position +78 is likely to adversely affect base pairing interactions within the bulge region to which Tat2 specifically binds (Peterlin & Jones, 1994). Nevertheless, as the mutation displayed by sequences from LTNP2 occurs within the second stem loop structure of the HIV-2 TAR element in the context of a fully functional stem-loop 1, the transcriptional effects of this point mutation are likely to be small. It is possible that the combined effects of the changes to the HIV-2 Inr and TAR sequence elements of LTNP2 result in the reduction in basal and Tat-induced transcriptional activity directed by the LTRs from this patient.

LTR sequences from LTNP3 were found to contain two base changes and a single base deletion within sequences flanking the TATA box motif. While very little data exists concerning the functional significance of these sequences within the HIV-2 LTR, several studies have demonstrated that site-directed mutagenesis of the nucleotides flanking the HIV-1 TATA box results in marked decreases in basal and Tat-induced gene expression (Garcia *et al.,* 1989, Berkout & Jeang, 1992, Ou *et al.,* 1994, Rittner *et al.,* 1995). It has been suggested that the sequences 5' and 3' to the HIV-1 TATA box might help regulate the binding affinity of members of pre-initiation complex (Berkout & Jeang, 1992, Ou *et al.,* 1994, Rittner *et al.,* 1995). In eukaryotic transcription, the sequences located nine base pairs upstream, and seven base pairs downstream of the eukaryotic TATA element (Lagrange *et al.,* 1998, Littlefield *et al.,* 1999) form the "IIB recognition element" (BRE), a bipartite motif bound by the basal transcription factor TFIIB. The role that this transcription factor plays in the formation of the transcription initiation complex is essential since, as well as binding to TFIID, TFIIB can also interact RNA polymerase II itself

(Tschochner *et al.*, 1992, Ha *et al.*, 1993). If, as with the other core elements of the HIV-2 LTR, the sequences flanking the HIV-2 TATA box function in an analogous manner to those within the HIV-1 and eukaryotic promoters, the mutations displayed by the sequences from LTNP3 might be expected to impair either the binding or regulation of transcription factors involved in transcription initiation such as TFIIB, and result in a decrease in the activity of LTRs from this patient.

Of the three RP patients, only LTR sequences from RP1 showed any significant nucleotide sequence variation within the functional sites of the LTR. LTR sequences from this patient displayed changes to the PuB1, *pets*, and peri-κB sites within the regulatory region of the HIV-2 LTR. Significantly, the mutation within the PuB1 site is located within the pentanucleotide core motif of the site. This motif acts as the recognition sequence for the ets-related transcription factor Elf-1, and changes to it have been shown to markedly reduce the transcriptional activity of the HIV-2 LTR *in vitro* (Leiden *et al.*, 1992, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hannibal *et al.*, 1994). In addition, mutagenic studies using *in vitro* assay systems have demonstrated the existence of a synergystic relationship between the *cis*-acting binding elements within the regulatory region of the HIV-2 LTR, since mutation of any of these sites results in a significant reduction in LTR function (Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994). Interestingly, in functional analyses the LTR sequences from RP1 were consistently found to be less active than the LTRs from both RP2 and RP3, patients whose LTRs contained consensus sequence motifs at most, or in the case of RP3, all regulatory, core and TAR elements of the LTR. It is possible that the mutations displayed within the PuB1 and *pets* sites of the LTRs from RP1 result in the disruption of transcription factor binding and synergystic interactions, and contribute to the comparatively reduced activity of the LTRs from this patient.

Of the six patients analysed, LTR sequences from RP3 were the only sequences to display 100% consensus sequence motifs at each of the single LTR functional sites within the LTR. Mutation rates, both general and G-to-A, at single LTR functional sites within the sequences from this patient were 0% in both cases. In addition, of the five patients that did bear mutations within one or more of the functional LTR sites, LTR sequences from RP2 showed the next lowest rates of mutation. Sequences from this patient showed a general mutation rate at single functional LTR sites of 1.1% and a G-to-A substitution rate at single functional LTR sites of 0%. Functional analyses of the basal and Tat-induced activities of the LTR sequences from RP2 and RP3 revealed that the sequences from both patients were consistently the most active of the LTRs tested.

In summary, while we were unable to detect a sequence motif that distinguished LTNP from RP LTRs, and failed to correlate LTR truncation, insertion or deletion with clinical long-term non-progression, we have shown the average general and specific G-to-A mutation rates to be higher within the single functional sites of LTNP LTRs when compared to RP LTR sequences. Moreover, of the functional sites within the HIV-2 LTR, LTR sequences from LTNP patients tended to display mutations within the highly conserved sites of the core and TAR regions of the LTR, while mutations within RP LTRs tended to occur within the less conserved transcription factor binding sites of the regulatory region of the LTR. Mutation of the highly conserved core transcription factor binding elements within the HIV-2 LTR has been shown to markedly reduce the basal activity of the HIV-2 LTR. Similarly, mutation of the stem-loop motifs within the TAR element of the HIV-2 LTR have been shown to significantly reduced the Tat-induced activity of the LTR. Therefore, it is possible that the general trend towards lower basal and Tat activities displayed by the LTNP LTRs within functional studies results from the higher average rate of mutation in the core and TAR transcription factor binding elements of the LTRs from these patients.

Significantly, the LTRs that had displayed exceptionally low (LTNP1) and notably high (RP3) basal and Tat-induced activities in the functional analyses were found to exhibit the highest and lowest rates of mutation respectively, throughout the LTR as a whole and at single functional LTR sites. Taken together our analyses indicate the existence of a relationship between the rate of mutation at single LTR functional sites, LTR activity, and clinical progression status. It is possible that the mutations within the LTNP LTR sequences have a greater effect upon function in Jurkat cells than THP-1 cells explaining the more pronounced difference in promoter activities observed between the two progression groups. However, studies performed with a greater number of clinically characterised patients in addition to SDM investigation of the mutations observed within the LTR sequences would be required to confirm such findings.

Chapter 8.
General Discussion and Conclusions.

8.1 General discussion

Since its initial isolation from patients with AIDS in Guinea Bissau and Cape Verde Islands in 1986 (Clavel *et al.,* 1986a), Human immunodeficiency virus type 2 (HIV-2) has been shown to be endemic in many West African countries including The Gambia, Guinea Bissau, Senegal, The Ivory Coast and Ghana (Decock & Brun Vezinet, 1989). Effective antiviral treatment and longitudinal clinical follow-up of the infected individuals within these populations while theoretically possible, is made practically impossible due to the economic cost of such schemes coupled with the rural and scattered nature of the populations requiring treatment. Therefore, the most likely means to prevent the spread of HIV-2 `infection and AIDS in these populations would be the provision of effective preventative and/or prophylactic vaccines.

Cohort studies in West Africa have demonstrated a markedly longer period of asymptomatic infection and prolonged survival in the majority of HIV-2-infected individuals compared to HIV-1 (Poulsen *et al.,* 1989, Wilkins *et al.,* 1993, Whittle *et al.,* 1994). Mortality rate ratios for HIV-2 are only 2:1 - 4:1 compared to uninfected controls (Ricard *et al.,* 1994, Poulsen *et al.,* 1997, Berry *et al.,* 2002, Schim van der Loeff *et al.,* 2002), and are ten-fold lower than mortality rate ratios observed in HIV-1 infection, where ratios exceed 20:1 (Muldor *et al.,* 1994, Morgan *et al.,* 1997, Nunn *et al.,* 1997, Todd *et al.,* 1997). Rates of both heterosexual and perinatal transmission are also significantly lower for HIV-2 than for HIV-1 (Adjorlolo-Johnson *et al.,* 1994, Kanki *et al.,* 1994, Cavaco-Silva *et al.,* 1998, O'Donovan *et al.,* 2000), this has resulted in a markedly different pattern of spread for the two infections. While the rapid spread of HIV-1 has lead to the emergence of a worldwide pandemic, HIV-2 infections have remained endemic and largely confined to the countries within West Africa (Remy, 1998). Taken together, HIV-2 is considered a less pathogenic virus than HIV-1. However, some cases of rapidly progressing HIV-2 AIDS following a more HIV-1-like disease course have been described both in The Gambia, and among West Africans resident in Europe (van der Ende *et al.*, 1996, Ariyoshi *et al.,* 1998).

The relationship between plasma viral loads and HIV pathogenesis has been widely documented; levels of virus found in the peripheral blood of HIV-1 and HIV-2-infected individuals have been shown to predict the rate of $CD4^+$ decline and are clearly associated with differences in rates of disease progression (Ho *et al.,* 1995, Wei *et al.,* 1995, Ariyoshi *et al.,* 2000, Berry *et al.,* 2002). Significantly, recent analyses

have demonstrated that HIV-1 and HIV-2 infections can be differentiated by RNA plasma viral load at the early stage of disease (Berry *et al.*, 1998, Andersson *et al.*, 2000). RNA virus is readily detectable in the plasma of the vast majority of HIV-1 infected individuals with a CD4% >28, while it is detected in only one third of matched HIV-2 infected subjects (Berry *et al.*, 1998). Following seroconversion, the plasma virus set point has been shown to be up to 28-fold lower in recent HIV-2 seroconverters when compared to recent HIV-1 seroconverters (Andersson *et al.*, 2000). Furthermore, the lower levels of plasma viremia that characterise the asymptomatic phase of HIV-2 infection can remain unchanged for periods in excess of five years (Berry *et al.*, 2002), or even persist until the symptomatic stages of disease (Andersson *et al.*, 2000). Significantly higher levels of plasma viremia are observed throughout the asymptomatic phase of HIV-1 infection and viral loads are only comparable between the two infections in the late stages of disease. In contrast to this, cellular proviral loads have been shown to be similar within HIV-1 and HIV-2 infected patients throughout the course of infection (Berry *et al.*, 1994, Ariyoshi *et al.*, 1996, Norrgren *et al.*, 1997b, Berry *et al.*, 1998, Sarr *et al.*, 1999, Popper *et al.*, 2000). Since the number of integrated proviral templates are similar between the two infections it has been suggested that the reduced number of circulating virions within the plasma of HIV-2 infected individuals result from a lower rate of virus production from these templates in HIV-2 infection compared to HIV-1 infection. Taken together, it is thought that the diminished pathogenicity of HIV-2 may be due to a lower rate of viral production within HIV-2 infected individuals.

Since the reduction in the number of CD4 cells and resultant immune deficiency is directly related to the level of viremia within an HIV-infected individual, the prolonged maintenance of a reduced level of circulating virus within an HIV-2 infected individual results in a slower destruction of their immune system via direct (cytotoxic) and indirect (apoptosis) mechanisms. A number of studies have demonstrated that the rate of total lymphocyte apoptosis is lower in HIV-2 infection compared with HIV-1 (Jaleco *et al.*, 1994, Michel *et al.*, 2000); similar observations have also been made in macaques infected with HIV-2 (Dittmer *et al.*, 1996). In addition, Sousa *et al* have recently demonstrated that higher proportions of CD4 and CD8 cells taken from HIV-2-infected individuals retain the ability to produce IL-2 when compared to the proportions taken from HIV-1 infected individuals with equivalent CD4 T cell counts (Sousa *et al.*, 2001).

In accordance with the cause-effect mathematical model proposed by Nowak for disease progression in HIV-1, a decreased rate of virus production effectively increases the time taken to reach the viral diversity threshold (Nowak *et al.*, 1991). Nowak suggests that throughout the asymptomatic phase of infection a

slow but steady increase occurs in the number of genetically distinct viral strains. Eventually a finite viral diversity threshold is reached, below which the immune system is able to regulate viral population growth, but above which the viral population induces the collapse of the CD4 T-cell population. The model predicts the appearance of more highly pathogenic strains once host immunological pressures have been removed (Nowak *et al.,* 1991). How the "virulent" viruses evolve or emerge in an HIV-infected host is not yet known. One suggestion is that such viruses while present throughout infection are effectively suppressed by antiviral activity, and re-emerge when antiviral activity is lost, or redevelop by genetic escape. Another possibility is that as immune antiviral activity reduces an increase in replicative cycles allows a relatively non-pathogenic infecting strain to mutate over time and acquire virulence. What is clear in HIV-2 infection is that the mechanism that results in the eventual increase in virus production and immune damage is more protracted within an HIV-2 infected individual than in an HIV-1 infected individual.

It has been suggested that the lower rate of virus production and the extended phase seen during most HIV-2 infections could be due to a more effective immunological response towards the viral population that reduces viral replication and antigenic diversity to undetectable levels (Gotch *et al.,* 1993, Rowland-Jones *et al.,* 1995, Bertoletti *et al.,* 1998, Whittle *et al.,* 1998), effectively increasing the time taken to reach the viral diversity threshold. While limited, some immunologic studies have demonstrated several features of the immune response to HIV-2 that do distinguish it from HIV-1 infection. In contrast to the majority of HIV-1 infected individuals (McAdam *et al.,* 1998), HIV-2-infected individuals often have a strong CTL response to HIV-2, which in some but not all cases show cross-clade reactivity to HIV-1 CTL epitopes (Bjorling *et al.,* 1993, Gotch *et al.,* 1993, Rowland-Jones *et al.,* 1995, Bertoletti *et al.,* 1998, Dorrell *et al.,* 2001). In addition, sera from some HIV-2-infected individuals have been found to contain broadly reactive neutralising antibodies showing extensive cross-reactivity to HIV-1, heterologous HIV-2, and SIV isolates (Robert Guroff *et al.,* 1992, Bjorling *et al.,* 1993). Nevertheless, these findings are not consistently identified in every HIV-2 long-term non-progressor. Moreover, several investigators have suggested that what appears to be a more effective immune response against HIV-2 may be a result of, rather than the cause of a viral population with a lower replicative capacity (Desrosiers, 1999, Berry *et al.,* 2002).

Much attention has been given to potential associations between the genetic background of the host and the rate of disease progression in HIV-1 and HIV-2 infection. Homozygosity for a 12-bp deletion within the coding sequence for chemokine receptor CCR-5 (CCR5-Delta32) has been associated with a reduced

susceptibility to HIV-1 infection (Dean *et al.*, 1996, Liu *et al.*, 1996). Similarly, polymorphic allelic variants of the chemokine receptor CCR-2, and the gene encoding stromal derived factor – 1 (SDF-1), the natural ligand for CXCR-4, are believed to be involved with delay in HIV-1 disease progression (Binley *et al.*, 1998, Winkler *et al.*, 1998, Carrington *et al.*, 1999). However, all such variants are rare in North, West, and Sub-Saharan African populations when compared to Caucasian populations making it unlikely that chemokine receptor variation plays a dominant mechanistic role in nonprogression among HIV-2-infected individuals.

The rate of disease progression in HIV-1 seropositive individuals has also been linked with the HLA class I genotype (Steel *et al.*, 1988, Kaslow *et al.*, 1990, Kaslow *et al.*, 1996, Magierowska *et al.*, 1999). Certain class I molecules, such as HLA B14 and HLA C8, are consistently linked with slow progression in HIV-1 infection while others, such as HLA A29, A11, B22 and the extended haplotype HLA A1/B8/DR3, are linked with rapid progression (Westby *et al.*, 1996, Hendel *et al.*, 1999). Studies on persistently seronegative and apparently uninfected prostitutes in The Gambia, whose initial exposure is most likely to have been to HIV-2, have suggested an association between the resistance of these individuals to HIV infection and their cross-reactive HLA-B35 restricted CTL response (Rowland-Jones *et al.*, 1995, Bertoletti *et al.*, 1998, Dorrell *et al.*, 2001). In addition, a recent study examining the relationship between HLA type, plasma viral load and survival among dually infected individuals revealed that patients with HLA type B58, or B35, B53, and B58 together, exhibited lower HIV-1 plasma viral loads and slower disease courses (Alabi *et al.*, 2003). However, larger prospective studies are required to confirm this finding. Therefore the relationships between HLA class I genotypes and disease progression rates remain complex, require extensive analyses with HIV-infected populations, and are unlikely to solely account for nonprogression in either HIV-1 or HIV-2 infection *per se*.

Alternatively, the extended asymptomatic phase seen during most HIV-2 infections could be virally determined by regions within the HIV-2 genome that relate directly to gene regulation and virus expression. By directing lower levels of virus production during infection virally encoded factors rather than immunological responses might determine a longer period of time ensuing before significant immunological damage occurs and/or the finite viral diversity threshold is reached. In turn, virally determined lower levels of virus production could be more effectively contained by immune surveillance. Cases of rapidly progressing disease might arise in this model due to the emergence of highly pathogenic strains that had acquired virulence through virally driven genetic diversity and selection of HIV-2

genomes with enhanced replicative capacities (Grassly *et al.*, 1998, Kimata *et al.*, 1999), or via initial infection with a highly pathogenic viral strain.

To date, few studies exist investigating the virally encoded determinants of HIV-2 pathogenesis (Albert *et al.*, 1990, Tong-Starksen *et al.*, 1990, Hannibal *et al.*, 1993, Akimoto *et al.*, 1998, Grassly *et al.*, 1998, Sekigawa *et al.*, 1998, Reeves *et al.*, 1999, Kokkotou *et al.*, 2000). The vast majority of data that have been collected thus far centres around the HIV-2 envelope and the differences that exist between its interactions with cellular receptors when compared to HIV-1. In general, HIV-2 strains use a wider range of coreceptors compared to HIV-1, potentially enabling infection of a broader range of cell-types *in vivo* (Guillon *et al.*, 1998, McNight *et al.*, 1998, Owen *et al.*, 1998, Unutmaz *et al.*, 1998, Reeves *et al.*, 1999, Liu *et al.*, 2000, Reeves & Doms, 2002). Moreover, in contrast to HIV-1, many primary HIV-2 isolates have been shown to infect cells independently of CD4 (Clapham *et al.*, 1992, Endres *et al.*, 1996, Reeves *et al.*, 1999, Simmons *et al.*, 2000, Lin *et al.*, 2001). This most recent discovery has been suggested to reflect a more 'open' HIV-2 Env glycoprotein conformation that might result in more effective immune recognition of HIV-2 and explain the prolonged survival of HIV-2-infected individuals when compared to HIV-1 (Reeves & Doms, 2002). However, the implications of CD4-independent virus infection for HIV-2 pathogenesis and its significance *in vivo* have yet to be explored, and may simply represent another mechanism by which the virus can establish a more successful virus-host relationship than that observed in HIV-1 infection.

The other region of the HIV-2 genome that has received investigative attention in the context of disease progression rate is the *nef* gene. In 1998, Switzer *et al* suggested a role for *nef* in the pathogenesis of HIV-2 infection (Switzer *et al.*, 1998). This study demonstrated that the prevalence of truncated Nef proteins was higher in HIV-2 asymptomatics when compared to the frequency found in HIV-1 equivalents (Switzer *et al.*, 1998). However, despite this finding Nef truncation was still only identified in 10% of the HIV-2-infected asymptomatic patients tested. More recently, a study performed by Padua and coworkers (Padua *et al.*, 2003) has identified a mutation within the tetra-proline motif of the Nef protein of HIV-2 asymptomatic individuals who also had low levels of peripheral viral RNA (11 of 26 asymptomatics - 42.3%). This mutation was not identified in any of the symptomatic individuals (11 of 11 symptomatics) within the study; the functional significance of this sequence change has yet to be determined. To date, a general explanation for the lower rate of virus production and the prolonged period of clinical long-term non-progression in HIV-2 infection has not been found.

Productive infection by HIV is dependent on continued activation of infected target cells (Finzi, 1998); lower production of virus in HIV-2-infected cells could reflect a lower activation state in those cells or that HIV-2 is less responsive to such activation. The LTR of both HIV-2 and HIV-1 regulates the expression of the proviral DNA in response to cellular transactivation signals. This region of the virus exerts therefore, a direct influence upon the rate of virus production. Sequence variations in the HIV LTR might affect the interaction with cellular factors and alter viral gene expression and the production of progeny virions in infected cells. Thus, natural LTR variations could have a significant impact on the efficiency of viral replication and the clinical course of HIV infection. The inherent pathogenicity of the HIV-1 LTR has been demonstrated *in vitro* whereby variant 5' LTR sequences have been shown to determine differences between the replicative capacities of particular HIV-1 (Golub *et al.,* 1990, Zhang *et al.,* 1997a) and SIV (Anderson & Clements, 1992) strains. Duplication of the complete SP1 region through prolonged culturing of an attenuated HIV-1 subtype B virus has been described, the LTR variant containing six SP1 sites was a stronger promoter and yielded a fitter virus (Berkout *et al.,* 1999). Similarly, uncultured HIV-1 LTR variants containing additional SP-1 binding sites have been shown to outgrow isogenic constructs with three SP-1 sites (Koken *et al.,* 1992).

More recently, functional distinctions in LTR architecture among HIV-1 subtypes have been identified, raising the possibility that regulatory divergence among the subtypes of HIV-1 has occurred (Montano *et al.,* 1997, Naghavi *et al.,* 1999, Hunt & Tiemessen, 2000, Jeeninga *et al.,* 2000, Hunt *et al.,* 2001, Rodenburg *et al.,* 2001, Novitsky *et al.,* 2002, Roof *et al.,* 2002, Scriba *et al.,* 2002). Transcriptional analyses performed by Jeeninga *et al* demonstrated that subtype E LTRs were consistently 2-3-fold more active at the basal level when compared to the LTRs of other subtypes A through G (Jeeninga *et al.,* 2000). When recombinant viruses were constructed; core promoter elements from subtype E were found to confer profound replicative advantage to subtypes with less active LTRs. A study by Verhoef *et al* has suggested that a single-nucleotide deletion in the upstream NF-κB element converts the site into a GABP binding element, contributing to the improved replication observed in subtype E viruses (Verhoef *et al.,* 1999). Similarly, within the same region of the LTR, several groups have demonstrated that a potential third NF-KB element within the subtype C LTR confers transcriptional and replicative advantage to subtype C virions (Hunt & Tiemessen, 2000, Hunt *et al.,* 2001, Rodenburg *et al.,* 2001, Novitsky *et al.,* 2002, Roof *et al.,* 2002, Scriba *et al.,* 2002). Since subtypes C and E have become the most prevalent and rapidly transmissible HIVs worldwide it has been suggested that this divergent transcriptional regulation may go some way towards explaining the observed differences in HIV transmission and pathogenesis.

Analyses of primary HIV-1 LTRs have revealed that basal and Tat-induced transcriptional activities vary considerably in infected individuals, and a variety of deletions and mutations have been detected (Delassus *et al.*, 1992, Koken *et al.*, 1992, Micheal *et al.*, 1994, Estable *et al.*, 1996 Kirchoff *et al.*, 1997, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Krebs *et al.*, 1998, Quiñones-Mateu *et al.*, 1998, Gomez-Roman *et al.*, 2000). Mutation at NF-κB and SP-1 sites and partial deletions of the HIV-1 LTR have been described in HIV-1 LTS with persistently undetectable virus in plasma (Fujii *et al.*, 1997, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Fang *et al.*, 2001) as has a higher G-to-A mutation rate in crucial sites of the LTR (Zhang *et al.*, 1997b, providing a possible genetic explanation for the low viral load and prolonged asymptomatic state in these individuals. Moreover, by analysing 20 complete HIV-1 RNA genomic sequences a recent study by Fang *et al* correlated the transition from long-term non-progressive to rapidly progressive infection with sequence changes to an SP-1 binding site and adjacent promoter within the HIV-1 LTR (Fang *et al.*, 2001). While correlations have been found between HIV-1 LTR polymorphism and the progression status of individual cases, to date, a relationship between HIV-1 LTR defects or genetic diversity and disease progression in general has yet to be established (Estable *et al.*, 1996, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Quiñones-Mateu *et al.*, 1998, Visco-Comandini *et al.*, 1999, Gómez-Román *et al.*, 2000).

In contrast to HIV-1, few functional studies exist concerning the HIV-2 LTR. Those that do exist have outlined significant differences between the structure and regulation of transcription from the two promoter regions (Markovitz *et al.*, 1990, Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995, Faulkner *et al.*, 2001). While the regulation of inducible promoter function in activated T cells is determined predominantly by the two NF-κB sites in the HIV-1 LTR (Nabel & Baltimore, 1987), HIV-2 LTR activation in T cells requires complex stereospecific interactions between the transcription factors that bind to at least four synergystic *cis*-acting binding elements (Markovitz *et al.*, 1992, Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994, Clark *et al.*, 1995). Studies have shown that the HIV-1 LTR is stimulated more effectively by TNF-α, a cytokine that acts primarily through NF-κB elements to activate virus replication in any T cell bearing its receptor. The HIV-2 LTR however, appears to be more effectively stimulated by the calcium-mediated signal transduction pathway activated by soluble anti-CD3 antibodies and TCR mediated antigen recognition (Hannibal *et al.*, 1993, Hilfinger *et al.*, 1993, Hannibal *et al.*, 1994). This method of activation is clonally specific and thus a more limited event. Since the regulation of the HIV-2 LTR significantly differs from the HIV-1 LTR, it has been suggested that the HIV-2 LTR could hold the key to

explaining the long clinical latency seen during HIV-2 infection, perhaps by contributing to a reduced rate in virus production or attenuated viral phenotype. Similarly, it is possible that a difference in HIV-2 LTR activity or regulation might determine the difference between the progression rates of the majority LTNP HIV-2-infected individuals and the cases of rapidly progressing HIV-2 AIDS, with enhanced or reduced promoter activity providing a possible mechanism for controlling the rate of antigenic diversification and consequent time taken to reach the viral diversity threshold.

The objectives of this project were two-fold. Firstly, to develop and optimise efficient PCR-amplification, cloning, transfection and detection procedures for the rapid and reproducible recovery and measurement of HIV-2 LTRs and their activities. This enabled a thorough examination of the biological significance of functional diversity of HIV-2 LTR genotypes with respect to viral pathogenicity and clinical disease. We hoped to elucidate functional differences between the HIV-2 LTRs cloned from patients and viral isolates with diverse *in vitro* phenotypes that may possibly lead to answers regarding the suggestion that the long clinical latency seen during HIV-2 infection is due to an 'attenuated' viral phenotype. Secondly, we examined the nucleotide sequence diversity of HIV-2 LTR genotypes in order to identify any correlation between LTR sequence polymorphism, LTR function, and disease progression rate, or outcome.

We have developed a limiting-dilution sensitive nested PCR that is capable of amplifying 764bp of the 3' HIV-2 LTR from a single molecule of HIV-2. This has enabled the amplification of the HIV-2 LTR from molecular clone, isolate, and high and low copy number patient-derived PBMC DNA. PCR amplicons generated by this procedure can be cloned into the experimental firefly luciferase reporter vector - pGL3E (Promega) using an optimised restriction digest cloning protocol that routinely achieves cloning efficiencies of up to 80%. HIV-2 LTR clones can then be co-transfected with an internal control vector into several different cell-lines using optimised co-transfection protocols that have been shown to be reliable and reproducible within each of the cell-lines used. The inclusion of an internal control vector, whose activity is used to normalise expression directed from the experimental reporter vector, has effectively eliminated inherent experimental variability's that arise from differences in transfection efficiencies and the number and health of cultured cells. This has enabled normalised luciferase activity values obtained from co-transfections carried out at different time points and in different cell-lines to be compared. Finally, the optimised dual-luciferase reporter detection assay has enabled the rapid and reproducible measurement of the activity of each cloned HIV-2 LTR in this study over a dynamic range spanning at least seven orders of magnitude. The efficiency and reproducibility of the nested-PCR,

cloning, transfection and reporter detection assay protocols developed in this thesis have also enabled HIV-2 LTR activity to be assessed in samples derived from the Caio cohort, Guinea Bissau (Okarofor,L *et al* unpublished). Furthermore, reaction conditions and techniques have been utilised to measure and compare the activities of LTRs derived from HIV-1-infected brain and lymph node tissue samples (Malik *et al* unpublished). The development of such a system therefore, constitutes a valuable addition to the currently available tools for those wishing to measure and compare promoter activity in functional characterisation studies.

PCR amplification of the HIV-2 LTR from molecular clone, isolate, and RP patient DNA was highly successful and readily achieved using a (non-optimal) nested PCR of high stringency. In contrast, amplification of the HIV-2 LTR from LTNP patient-derived PBMC DNA proved to be more difficult using the high stringency PCR, and was only achieved by doubling the initial first round DNA concentration and reducing the annealing temperature of the second round PCR from 54 – 50°C. Previous studies concerned with PCR amplification from HIV-2 infected PBMC material have experienced similar difficulties when trying to amplify from asymptomatic patient samples. A study by Simon *et al* found that the ability to detect the HIV-2 genome by PCR was a reflection of the level of virus within the patient, and correlated with CD4 count. Patients exhibiting $CD4^+$ counts of >500 x 10^6/l were less likely to carry detectable levels of HIV-2 DNA than patients exhibiting $CD4^+$ counts of <200 x 10^6/l. A characteristic of the Gambian LTNP patients is very low to undetectable plasma viral loads (1 copy 1000 to 1 copy in 10,000 $CD4^+$ cells); in contrast, rapidly progressing Gambian patients exhibit high plasma viral loads (up to 1 copy in 10 $CD4^+$ cells). It is likely therefore, that amplification from the three LTNP samples initially failed, in part, because of the very low levels of proviral copies within the PBMC derived DNA. The limiting dilution sensitivity of the optimised HIV-2 LTR nested PCR has enabled the amplification of the 764bp LTR region from all available DNA sources.

Various cloning and transfection techniques are currently available for use in molecular characterisation studies however; each must be optimised for the particular system within which they will be applied. By increasing the vector to inset ratio within the cloning reaction we were able to routinely achieve 70 – 80% cloning efficiencies. Similarly, increasing the ratio at which the luciferase vectors were cotransfected reduced *trans* effects between them in cotransfection reactions. Extensive optimisation of the transfection techniques available for each cell-line used in this study enabled detectable and reproducible firefly and *Renilla* luciferase measurements to be made, and ensured that normalised luciferase activity values obtained from co-transfections carried out at different time points could be compared. Washing

transfected cells 2 hours post-transfection was not found to significantly increase normalised luciferase values or decrease growth inhibition and cell death. However, cell density was found to significantly alter reporter gene expression, with lower levels of gene expression being measured at lower cell densities. Optimal reporter gene expression was obtained when 2 x 10^6 Jurkat or THP-1 cells were transfected.

Under optimal conditions, analyses of the reproducibility of the dual luciferase detection assay and reproducibility of co-transfection revealed no significant differences between the levels of normalised luciferase activity taken from each well assayed in duplicate (R^2 = 0.9888), or from each well of a duplicate co-transfection (R^2 = 0.9848). Furthermore, very little variation was detected between the normalised luciferase values obtained from identical co-transfections performed one week apart (R^2 = 0.8796). The sensitivity and repeatability of the optimised protocols within the dual luciferase detection assay system has ensured that reporter protein expression directed by promoters of varying activities can be measured in singlicate, accurately and reproducibly.

Preliminary studies into the functional characterisation of the HIV-2 LTR were performed on a series of Gambian isolates (CBL20-24). CBL-20 and CBL-21 were derived from AIDS/ARC patients whereas CBL-23 and CBL-24 were derived from asymptomatic individuals. Analysis of these materials was instructive both in terms of the work-up of the technology and in generating preliminary data to enable the study to be extended to primary uncultured PBMC samples. However, the culturing of viruses *in vitro* may not provide a completely accurate representation of the viral population originally circulating in the patient when the isolate was first made, thus analyses of this kind require this observation to be taken into account.

Functional characterisation of the LTRs derived from the CBL series of HIV-2 isolates by the dual luciferase assay system, identified significant variation between the activities of LTRs cloned from the same and from the different CBL isolates when tested in both the Jurkat and THP-1 cell-lines (representing T-cell and macrophage-like cell lines respectively). This variation appeared to be reproducible, and is at least in part likely to be a reflection of viral polymorphism within each HIV-2 isolate. It has been known for some time that HIV exists as a quasispecies or 'swarm' of viruses *in vivo*, where cellular proviruses are present as a library of similar, though not identical, sequences (Goodenow *et al.*, 1989). The demonstration of the transcriptional heterogeneity within and between the HIV-2 LTR clones from the five HIV-2 isolates not only reflects this finding but demonstrates that functional

heterogeneity can persist despite the artificial selection conditions of culturing. This in turn allowed instructive data to be obtained and encouraged further study of viral populations *in vivo.*

Interestingly, examination of the LTR activities at both the basal and Tat-induced levels revealed a non-significant trend towards higher basal and Tat-induced activities in Jurkat cells for the LTRs cloned from the more rapidly replicating CBL isolates, CBL-20 and CBL-21. The mean Tat-induced activity of the LTRs from CBL-20 was found to be statistically higher than two of the less rapidly replicating isolates (CBL-22 and CBL-23) in Jurkat cells. In addition, the intraisolate range of LTR activities from both CBL-20 and CBL-21 were greater than the intraisolate activity ranges of LTRs from the other CBL isolates in the Jurkat cell-line.

Differences of only two- and three-fold in the basal activity of the HIV-1 LTR have previously been shown to determine the different replicative capacities of two closely related HIV-1 isolates (Golub *et al.,* 1990). Therefore, it is possible that the range of basal and Tat-induced LTR activities of CBL-20 and CBL-21 in Jurkat cells reflect the phenotypes of these isolates and broadly correlate with the progression rates of the patients from whom these isolates were derived. However, while this non-significant trend in LTR activity was observed in Jurkat cells, there was no overall association between either basal or Tat-induced LTR activity and viral phenotype *in vitro.* This finding is in agreement with two studies performed with the cultured LTR quasispecies of slowly and rapidly replicating HIV-1 isolates (McNearney *et al.,* 1995, Simm *et al.,* 1996) which failed to distinguish between the HIV-1 isolates on the basis of transcriptional activity. There may be a number of explanations for these findings.

In some respects it is not surprising that higher levels of LTR activity were observed in patients with more advanced disease, although whether a more active LTR is having a causal effect on disease progression or merely a reflection of it is not known. One possible interpretation is that the clinical status of the HIV-2 infected individuals from whom the CBL isolates were derived is not determined solely by the LTR. However, this does not preclude the possibility that the non-significant trend in activities observed in the Jurkat cell-line might prove significant in different cell-types *in vivo* and may therefore contribute in some way to determining their differing *in vitro* phenotypes. The differing pattern of LTR activities observed in the THP-1 cell-line gave the first indication in this thesis that the activity of the HIV-2 LTR may be related to cell-type. It is also possible that functional differences that may have existed between the LTRs of the CBL isolates before prolonged passage in culture may have been distorted due to the highly selective culture environment. The broader criticism that the use of *in vitro*

viral isolates is unrepresentative of the viral swarm *in vivo* precludes over-interpretation of these findings.

Based on the data generated in the CBL series study we were encouraged to go forward and investigate LTR activity with respect to disease progression using PBMC derived DNA that had not been subjected to the selective pressures of *in vitro* culturing. The basal and Tat-induced activities of LTRs amplified from the PBMC-derived DNA of two clinically contrasting groups of HIV-2 seropositive patients were compared in both the Jurkat and the THP-1 cell-line. Significantly, some level of transcriptional activity was detected from the HIV-2 LTRs taken from each of the three LTNP and three RP patients within the analyses. This finding, along with the observations made with the LTRs of CBL isolates 23 and 24 indicates that the prolonged freedom from clinical illness experienced by the long-term non-progressing and asymptomatic HIV-2 patients does not result from the presence of a transcriptionally attenuated quasispecies within these patients.

Similar to findings from several HIV-1 LTR studies (Fujii *et al.*, 1997, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b), we did identify one individual whose clinical long-term non-progression is most likely explained by defective LTR sequences. LTRs cloned from patient LTNP1 were found to direct consistently low basal activities which were not increased by any detectable amount in the presence of the HIV-2 Tat protein. In both Jurkat and THP-1 cell lines the activities of LTRs from the remaining five LTNP and RP patients were found to increase over basal levels by at least 15-fold. These data demonstrate that the function of the LTRs from LTNP1 was severely impaired. Cloning of LTRs from this and other patients was performed on at least two separate occasions by bulk PCR using fresh aliquots of PBMC derived DNA. Therefore, it is likely that the LTRs used to represent LTNP1 in these analyses reflect the circulating LTR genotypes within this individual, rather than LTRs containing errors introduced during the process of PCR amplification or a minority LTR species. When compared to the HIV-2 ROD LTR, sequences from LTNP1 contained multiple G-to-A hypermutations within the NF-κB, SP-1, and TAR binding elements of the LTR. Such hypermutations are expected to have a significant impact on the function of the LTR since they are within critical core promoter binding motifs. Mutations of this kind are predicted to severely disrupt the binding of the transcription factors NF- κB and SP1 and the cellular cofactor for Tat – CycT1 respectively (Arya & Gallo, 1988, Markovitz *et al.*, 1990, Tong-Starksen *et al.*, 1990, Markovitz *et al.*, 1992, Pagtakhan & Tong-Starksen, 1997 Richter *et al.*, 2002a, Richter *et al.*, 2002b). In 1997, during the genotypic and phenotypic characterisation of LTR sequences from LTS of HIV-1 infection, Zhang and co-workers identified LTRs from a LTS patient (Pt5) that

exhibited similar G-to-A hypermutations to those identified in the LTR sequences of LTNP1 (Zhang *et al.*, 1997b). Analysis of LTR function by luciferase reporter assay system revealed that the activities of the LTR sequences from Pt 5 were severely blunted. Similarly, when whole infectious virus studies were performed, no detectable level of viral transcription was found in PBMC infected by proviral constructs containing Pt 5 LTR sequences. From these data Zhang *et al* concluded that it was likely that the viruses in Pt 5 were some-what attenuated, and that Pt 5's functionally defective LTR provided some explanation for his prolonged period of well-being (Zhang *et al.*, 1997b). While we have not performed whole infectious viruses studies with the LTR sequences from LTNP1, we have observed similar structural and functional finding to those identified by Zhang and co-workers. It is probable therefore, that the undetectable plasma viral loads and prolonged freedom from clinical illness exhibited by LTNP1 are in part explained by attenuation of LTR function.

Regarding LTR function in the context of progression groups, significant variation was observed between the basal and Tat-induced activities of the LTNP and RP LTRs. We have shown that LTRs cloned directly from the PBMCs of patients exhibiting long-term non-progression direct markedly lower levels of activity on average than the activities of LTRs from patients exhibiting unusually rapid progression to disease. While no direct correlation between LTR activity and disease progression status has been identified, a non-significant trend towards lower LTNP LTR activity has been observed in both Jurkat and THP-1 cell-lines. This observation was most evident at the Tat-induced level of activity within Jurkat cells. LTNP LTRs were found to be, on average, three times less active than the RP LTRs. Moreover, when analysed individually, the average LTR activities of two of the three LTNP patients were found to be statistically significantly lower than the LTRs of the two most active RP patients. These findings differ from a number of HIV-1 LTR studies that have found no association between LTR activity and progression status when LTR function in long-term survivors and typical progressors of HIV-1 infection was compared (Estable *et al.*, 1996, Rousseau *et al.*, 1997, Zhang *et al.*, 1997b, Quiñones-Mateu *et al.*, 1998, Gómez-Román *et al.*, 2000). Most were performed using comparable reporter assays and similar cell-lines, ruling out the possibility that the differing conclusions result from different techniques. Rather, the differences noted within the analyses presented in this thesis are genuine and reproducible and likely represent a functional difference between the activities of the promoter regions of the virions circulating within the LTNP and RP patients in these analyses. As stated at the outset of this thesis, this is the first functional comparison of naturally occurring HIV-2 LTR genotypes in the context of disease progression and pathogenesis and the results are therefore novel.

An explanation for the non-significant trend observed in LTR activities may lie in the differences noted between the rates of mutation within the LTNP and RP LTRs. The nucleotide sequence analyses performed upon the HIV-2 LTR sequences in this study suggest that there may be a correlation, albeit non-significant, between the rate of mutation at single functional sites, HIV-2 LTR activity, and clinical progression status. While the general mutation rate throughout the LTR as a whole was found to be similar between the two progression groups, the findings outlined in chapter 7 demonstrated that LTR sequences from the LTNP group of patients had on average 2.5 times more general mutations at single functional sites than the RP group of LTR sequences. Moreover, on average, LTNP LTRs contained twice the number of G-to-A mutations throughout the entire LTR sequence, and five times the number of G-to-A mutations at single functional sites when compared to RP LTR sequences.

Though in the absence of site-directed mutagenesis (SDM) analysis it is difficult to say what the functional effect of the higher rate of mutation at single functional LTR sites in LTNP LTRs would be, it is likely to result in some kind of transcriptional disadvantage. Significantly, the two patients that exhibited the highest rates of mutation were the LTNPs that directed the lowest LTR activities. Conversely, the two patients that exhibited the lowest rates of mutation were the RPs that directed the highest LTR activities. Most importantly, sequence changes within the LTNP LTRs were located almost entirely within the core transcription factor binding sites of the HIV-2 promoter, while RP LTRs displayed 100% consensus sequences for each of the binding sites within this region. Any mutations that were noted within the binding sites of the RP LTRs were located within the less conserved binding sites of the modulatory region of the LTR. It is therefore, not surprising that the LTNP LTR sequences tended to direct lower basal and Tat-induced activities than the RP LTRs. Changes to the functional sites within the core promoter region are predicted to have a significantly greater impact upon LTR function than changes to those elements situated further upstream since they are directly involved in binding the transcription factors involved in the assembly of the PIC (Arya & Gallo, 1988, Markovitz *et al.,* 1990, Tong-Starksen *et al.,* 1990, Markovitz *et al.,* 1992). In accordance with this, LTRs from RP1 that contained changes within the modulatory binding sites were shown to retain a greater ability to respond to the Tat protein than the LTRs from the three LTNP patients that contained changes within core promoter binding sites, despite the LTRs from RP1 having relatively low basal activities. It is possible that the lower basal and Tat-induced activity of the LTNP LTRs results from a reduction in the rate of pre-initiation complex (PIC) assembly caused by the effects of sequence changes within the binding elements located in the core region of the promoter. Mutations such as G-to-A substitutions result in completely

different functional groups being presented for interaction with the transcription factor binding proteins, and cause a reduction in the binding affinity and recruitment of those factors. Nevertheless, it is notable that these analyses did not identify a structural motif that distinguished the LTRs from LTNP and RP patients.

Numerous HIV-1 reports have correlated genetic diversity with disease progression, with greater sequence diversity being noted in viral populations of individuals who remain asymptomatic for a prolonged period of time compared to individuals showing more rapid progression to disease (Delwart *et al.*, 1994, Lukashov *et al.*, 1995, Wolinsky *et al.*, 1996). This observation is thought to be determined by constraints placed upon the viral genome by the immune system. The higher rate of mutation at single functional sites observed within the LTRs from the LTNP patients in these analyses appear in part to reflect this finding. However, while a number of individual cases of long-term non-progression in HIV-1 infection have been correlated with a higher rate of mutation within the promoter region (Fujii *et al.*, 1997, Zhang *et al.*, 1997b, Fang *et al.*, 2001), no definite correlation has been made between the rate of mutation within the LTR region of the viral genome and the rate of disease progression in HIV-1 infection *per se*. Therefore, analyses of a larger number of HIV-2 infected individuals would be required to confirm any association between mutation rates in the LTR, transcriptional activity, and progression status in HIV-2 infection.

Interestingly, the significance of the trends and associations noted between mutation rates, LTR activity, and progression status appear to be related to cell-type, with the greatest differences between the LTR activities of the two progression groups being noted in the T-cell-like cell-line. The LTR has been shown to be important in determining the cell tropism of a number of virus from the *Lentivirus* genus including, equine infectious anaemia virus (EIAV) (Carvalho *et al.*, 1993, Maury *et al.*, 1997, Maury, 1998, Payne *et al.*, 1999, Maury *et al.*, 2000, Hines *et al.*, 2004), feline immunodeficiency virus (FIV) (Poeschla *et al.*, 1998), and maedi-visna virus (MMV) (Andresdottir *et al.*, 1998, Agnarsdottir *et al.*, 2000). Moreover, cell-type specific differences in the activity of LTRs have been widely documented in HIV-1 analyses (Corboy *et al.*, 1992, Chang *et al.*, 1993, Moses *et al.*, 1994, Ait-Khaled *et al.*, 1995, Canonne-Hergaux *et al.*, 1995, Estable *et al.*, 1996, Corby & Garl, 1997, Henderson & Calame, 1997, Krebs *et al.*, 1997, Jeeninga *et al.*, 2000, McAllister, *et al.*, 2000). It has been demonstrated that some *in vivo*-selected LTRs can confer transcription and replication advantages in specific cell types leading to suggestions that several non-envelope factors, including the HIV LTR, may play an important role in determining cellular tropism in HIV-1 infection (Chang *et al.*, 1993, Ait-Khaled *et al.*, 1995, Henderson & Calame, 1997,

Krebs *et al.*, 1997, Jeeninga *et al.*, 2000, McAllister, *et al.*, 2000). In accordance with this, Reed-Inderbitzin *et al* recently demonstrated that it was possible to alter the cell tropism of HIV-1 by making substitutions within the enhancer region of the LTR, indicating for the first time that macrophage specificity of productive HIV replication could be mediated by LTR sequences (Reed-Inderbitzin *et al.*, 2003). Since the cellular tropism of HIV may play an important role in disease progression, with primary stages of infection being associated with viral species exhibiting macrophage-tropic non-syncytium-inducing replication characteristics and clinical disease progression being associated with the development of viral species exhibiting T-cell-tropic, syncytium-inducing phenotypes, it may be of significance that the RP LTRs in our analyses were found to direct significantly higher levels of transcription within the Jurkat T-cell-like cell-line than the LTRs from the LTNP patients. If in comparison to RP patients, long-term non-progressors of HIV-2 infection possess a viral population that directs viral gene expression and subsequent virus production at a lower level within cells that the virus preferentially replicates, the result could be a slower rate of immune damage via direct (cytotoxic) and indirect (apoptotic) mechanisms, and a slower rate of disease progression. In this sense it could be suggested that in the vast majority of cases HIV-2 is behaving more akin to a classical lentivirus (Gorrell *et al.*, 1992, Maury, 1994) than HIV-1, tending to be macrophage-tropic, or at least not showing the rapid/high replication properties within T cells demonstrated by HIV-1. Conversely if the rapid progressors of HIV-2 infection possess LTRs that provide transcriptional and thus replicative advantage within T cells in an 'HIV-1-like' manner, this may in turn influence the rate of virus production, rate of viral diversity, rate of immune damage and finally rate of progression to disease. Importantly, the findings derived from both the CBL isolate and Gambian patient cohort analyses indicate that the HIV-2 LTR may be involved in determining cell-type-specific transcription and replication. It is possible that the mutations that occur within the core transcription binding sites of the LTNP LTRs are tolerated less well in Jurkat cells than in THP-1s, leading to the more pronounced difference in LTR activity observed between the progression two groups in Jurkat cells.

It cannot be ignored that the data set within which these observations have been made is small (n=6), therefore analyses of HIV-2 LTR sequence and function within larger populations and/or cohorts would be required to investigate fully the role that the HIV-2 LTR plays in the pathogenesis of HIV-2 infection. The study number may provide one explanation as to why the findings presented here contradict those reported in the analyses of HIV-1 LTR function in LTS of HIV-1 infection. Numbers in the HIV-1 studies have ranged from eight patients (Rousseau *et al.*, 1997, Zhang *et al.*, 1997b) up to 42 patients

(Estable *et al.*, 1996, Quiñones-Mateu *et al.*, 1998), it is possible that the statistical significance of the findings presented in this thesis would be reduced were the numbers of patients in the study to be increased. Nevertheless, the differences that were noted between the progression groups as a whole in terms of LTR function and structure were in many cases statistically significant when the data from the fifteen LTR sequences as opposed to the three patient averages from each group were compared. Therefore, it maybe that increasing the patient number in the study would strengthen the significance of these findings.

In making the case for the causal involvement of the HIV-2 LTR in determining pathogenesis, the pathogenic influence of the LTR has been demonstrated in a number of other retroviral infections. For example, both the latency period of disease induction and the pathogenic/oncogenic potential of the avian leukosis virus group have been found to correlate with LTR transcriptional activity (Tsichilis & Lazo, 1991). In addition, acquisition of virulence by some strains of simian immunodeficiency virus has been associated with changes to viral genetic elements including the LTR region (Novembre *et al.*, 1993, Flaherty *et al.*, 1997, Karlsson *et al.*, 1997). Similarly, a recent study by Grassley *et al* revealed a significant correlation between the death of HIV-2 infected individuals and viral genetic identity when LTR, *gag*, and *pol* sequences were jointly aligned (Grassley *et al.*, 1998). Finally, the analyses performed by Fang *et al* in 2001 indicated that progression to symptomatic disease following a prolonged period of asymptomatic HIV-1 infection could be associated with sequence changes to the HIV LTR region (Fang *et al.*, 2001). It is possible therefore, that the differences in promoter activity that have been noted between the LTNP and RP patients in the analyses could reflect a mechanism by which, in the majority of infections, the HIV-2 LTR helps to determine the establishment of a more low-key and sustainable virus-host relationship than observed in HIV-1 or rapidly progressing HIV-2 infection. Analyses performed by Jordan *et al* have indicated that the basal activity of the HIV-1 LTR varies considerably depending on the integration site (Jordan *et al.*, 2001). It is possible that the site of integration within host cell DNA could further compound LTR activity in LTNPs that may already be compromised.

What is apparent is the need for further investigation of the pathogenic significance of the differences in activity observed between the LTNP and RP LTRs *in vivo*. It would be interesting to determine whether the observed transcriptional differences translate into replicative differences between proviral constructs containing the LTNP and RP LTRs analysed within this study. Transcriptional differences noted between HIV-1 LTRs in reporter gene analyses performed by Jeeninga *et al* were found to confer significant and comparative replicative differences to constructs containing the promoter regions under investigation

(Jeeninga *et al.*, 2000). Similarly, Zhang and coworkers demonstrated that the levels of transcription directed by HIV-1 LTRs in the context of whole infectious virus studies were remarkably consistent with the results obtained using the luciferase gene alone (Zhang *et al.*, 1997). From this it is possible to assume that the modest average two- (THP-1 cells) to three-fold (Jurkat cells) difference observed between the Tat-induced activities of the LTNP and RP LTRs could also reflect and confer significant differences to the replicative capacities of the viral populations circulating within the infected individuals from which they were taken. Similarly, this may in part influence the difference between the levels of viremia observed in the two groups of patients, with LTNP LTRs of lower activity directing lower levels of gene expression and virus production than RP LTRs of higher activities directing higher rates of gene expression and virion production. Since the level of virus production within HIV infection is thought to be related to the rate of disease progression and pathogenesis it is possible that the difference observed between the activities of the LTRs from the LTNP and RP patients are related to the differences between the rates of disease progression shown by the two groups.

Taken together the data presented in this thesis have indicated the existence of a non-significant trend towards lower LTR activity in HIV-2-infected patients exhibiting long-term non-progression. Since the observed trends are non-significant and we have failed to distinguish between LTNP and RP patients based upon LTR activity or structural LTR motifs alone it is unlikely that the HIV-2 LTR is the sole determinant of HIV-2 pathogenicity. Rather, the findings presented here indicate that the HIV-2 LTR may play a contributory role in determining pathogenesis via its influence upon gene expression and virus production, and have indicated that this role may be cell-type-specific. The data do not rule out the possibility that the proviral DNA in the LTNP and RP patients contain deletions or mutations in regions of the LTR that have yet to be identified as important for replication. Neither do they preclude the possibility that the differences in activities between the two groups that have been noted in the cell-lines tested could be greater within other cell-lines. However, they do suggest that long-term non-progression in HIV-2 infection is a multifactorial process almost definitely involving both host and viral determinants. In this vein, the analyses presented here does not exclude the possibility that other regions of the HIV-2 genome within each of the infected patients are playing a more significant role in the determination of rate of disease progression within each individual, or that the differences in LTR activity observed in this study merely reflect the appearance of more highly pathogenic strains during the development of symptomatic disease. Indeed, caution needs to be taken in drawing conclusions from the presented data as only the proviral population was sampled to compare LTR populations in LTNP and RP groups.

Archival populations may be misleading, so for a more accurate reflection of the impact that LTR sequences variation could have on influencing HIV-2 biology LTR sequences in plasma virus populations should also be studied. However, one of the confounding issues here concerns the very low levels of HIV-2 RNA circulating in the plasma of a true long-term non-progressor of HIV-2 infection. If ultra-sensitive RT-PCR methodologies were to be employed for this analysis high sample input volumes would be required and availability of such material is a significant problem. Hence, the low level of viral replication in true HIV-2 LTNPs presents considerable technical obstacles if one wishes to study the properties of virus populations which may be regarded as low-level replication competent viruses. How the sequence of these virus populations differs from proviral species in HIV-2 infections is not known but would be an interesting lead to follow up with the methodologies described. These issues underline the need for further research into the pathogenic determinants of HIV-2 infection. In conclusion, the findings presented in this thesis have provided a unique insight into the genotypic and phenotypic characteristics of naturally occurring HIV-2 LTRs and may be important for the future development of treatments or vaccines for both HIV-2 and HIV-1 infection.

References.

Aaby, P., Ariyoshi, K., Buckner, M., Jensen, H., Berry, N., Wilkins, A., Richard, D., Larsen, O., Dias, F., Melbye, M. & Whittle, H. (1996). Age of wife as a major determinant of male-to-female transmission of HIV-2 infection: a community study from rural West Africa. *AIDS* **10,** 1585-1590.

Adhya, S. & Gottesman, M. (1982). Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell* **29,** 939-944.

Adjorlolo-Johnson, G., De Cock, K.M., Ekpini, E., Vetter, K.M., Sibailly, T., Brattegaard, K., Yavo, D., Doorly, R., Whitaker, J.P., Kestens, L., Ou, C.Y., George, J.R. & Gayle, H. (1994). Prospective comparison of mother-to-child transmission of HIV-1 and HIV-2 in Abidjan, Ivory Coast. *J.A.M.A* **272,** 462-466.

Agnarsdottir, G., Thorsteinsdottir, H., Oskarsson, T., Matthiasdottir, S., St Haflidadottir, B., Andresson, O.S. & Andresdottir, V. (2000). The long terminal repeat is a determinant of cell tropism of maedi-visna virus. *J. Gen Virol.* **81,** 1901-1905.

Ait-Khaled, M., McLaughlin, J.E., Johnson, M.A. & Emery, V.C. (1995). Distinct HIV-1 long terminal repeat quasispecies present in nervous tissues compared to that in lung, blood and lymphoid tissues of an AIDS patient. *AIDS* **9,** 675-683.

Akimoto, H., Kaneko, H., Sekigawa, I., Hashimoto, H., Kaneko, Y. & Yamamoto, N. (1998). Binding of HIV-2 envelope glycoprotein to CD8 molecules and related chemokine production. *Immunology* **95,** 214-218.

Alabi, A.S., Jaffar, S., Ariyoshi, K., Blanchard, T., Schim Van Der Loeff, M., Awasana, A.A., Corrah, T., Sabally, S., Sarge-Njie, R., Cham-Jallow, F., Jaye, A., Berry, N. & Whittle, H. (2003). Plasma viral load, CD4 cell percentage, HLA and survival of HIV-1, HIV-2, and dually infected Gambian patients. *A.I.D.S.* **17,** 1513-1520.

Alam, J. & Cook, J.L. (1990). Reporter genes: application to the study of mammalian gene transcription. *Anal. Biochem.* **188,** 245-254.

Albert, J., Bottiger, B., Biberfeld, G. & Fenyo, E.M. (1989). Replicative and cytopathic characteristics of HIV-2 and severity of infection. *Lancet* **1,** 852-853.

Albert, J., Naucler, A., Bottiger, B., Broliden, P.A., Albino, P., Ouattara, S.A., Bjorkegren, C., Valentin, A., Biberfeld, G. & Fenyo, E.M. (1990). Replicative capacity of HIV-2, like HIV-1, correlates with severity of immunodeficiency. *AIDS* **4,** 291-295.

Albini, A., Barillari, G., Benelli, R., Gallo, R.C. & Ensoli, B. (1995). Angiogenic properties of human immunodeficiency virus type 1 Tat protein. *Proc. Natl. Acad. Sci. USA.* 92(11):4838-42.

Alonso, A., Cujec, T.P. & Peterlin, B.M. (1994). Effects of human chromosome 12 on interactions between Tat and TAR of human immunodeficiency virus type 1. *J. Virol.* **68,** 6505-6513.

Anderson, M.G. & Clements, J.E. (1992). Two strains of SIVmac show differential transactivation mediated by sequences in the promoter. *Virology* **191,** 559-568.

Andersson, S., Norrgren, H., da Silva, Z., Biague, A., Bamba, S., Kwok, S., Christopherson, C., Biberfeld, G. & Albert, J. (2000). Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. *Arch. Intern. Med.* **160,** 3286-3293.

Andreasson, P.A., Dias, F., Naucler, A., Andersson, S. & Biberfeld, G. (1993). A prospective study of vertical transmission of HIV-2 in Bissau, Guinea-Bissau. *AIDS* **7,** 989-993.

Andresdottir, V., Tang, X., Agnarsdottir, G., Andresson, O.S., Georgsson, G., Skraban, R., Torsteinsdottir, S., Rafnar, B., Benediktsdottir, E., Matthiasdottir, S., Arnadottir, S., Hognadottir, S., Palsson, P.A. & Petursson, G. (1998). Biological and genetic differences between lung- and brain-derived isolates of maedi-visna virus. *Virus Genes* **16,** 281-293.

Ariyoshi, K., Cham, F., Berry, N., Jaffar, S., Sabally, S., Corrah, T. & Whittle, H. (1995). HIV-2-specific cytotoxic T-lymphocyte activity is inversely related to proviral load. *AIDS* **9,** 555-559.

Ariyoshi, K., Berry, N., Wilkins, A., Ricard, D., Aaby, P., Naucler, A., Ngom, P.T., Jobe, O., Jaffar, S., Dias, F., Tedder, R.S. & Whittle, H. (1996). A community-based study of human immunodeficiency virus type 2 provirus load in rural village in West Africa. *J. Infect. Dis.* **173,** 245-248.

Ariyoshi, K., Schim van der Loeff, M., Cook, P., Whitby, D., Corrah, T., Jaffar, S., Cham, F., Sabally, S., O'Donovan, D., Weiss, R.A., Schulz, T.F. & Whittle, H. (1998). Kaposi's sarcoma in the Gambia, West Africa is less frequent in human immunodeficiency virus type 2 than in human immunodeficiency virus type 1 infection despite a high prevalence of human herpesvirus 8. *J. Hum. Virol.* **1,** 193-199.

Ariyoshi, K., Jaffar, S., Alabi, A.S., Berry, N., Van der Loeff, M.S., Sabally, S., N'Gom, P.T., Corrah, T., Tedder, R. & Whittle, H. (2000). Plasma RNA viral load predicts the rate of CD4 T cell decline and death in HIV-2-infected patients in West Africa. *AIDS* **14,** 339-44.

Arthos, J., Deen, K.C., Chaikin, M.A., Fornwald, J.A., Sathe, G., Sattentau, Q.J., Clapham, P.R., Weiss, R.A., McDougal, J.S., Pietropaolo, C., et al. (1989). Identification of the residues in human CD4 critical for the binding of HIV. *Cell* **57,** 469-481.

Arya, S.K., Beaver, B., Jagodzinski, L., Ensoli, B., Kanki, P.J., Albert, J., Fenyo, E.M., Biberfeld, G., Zagury, J.F., Laure, F., et al. (1987). New human and simian HIV-related retroviruses possess functional transactivator (tat) gene. *Nature* **328,** 548-550.

Arya, S.K. & Gallo, R.C. (1988). Human immunodeficiency virus type 2 long terminal repeat: analysis of regulatory elements. *Proc. Natl. Acad. Sci. USA.* **85,** 9753-9757.

Arya, S.K. (1991). Human immunodeficiency virus type 2 (HIV-2) gene expression: downmodulation by sequence elements downstream of the transcriptional initiation site. *AIDS Res. Hum. Retroviruses* **7,** 1007-1014.

Arya, S.K. (1993). Human immunodeficiency virus type 2 (HIV-2) trans-activator (Tat): functional domains and the search for trans-dominant negative mutants. *AIDS Res. Hum. Retroviruses* **9,** 839-848.

Ayanian, J.Z., Maguire, J.H., Marlink, R.G., Essex, M. & Kanki, P.J. (1989). HIV-2 infection in the United States. *N. Engl. J. Med.* **320,** 1422-1423.

Baeuerle, P.A. & Baltimore, D. (1988). I kappa B: a specific inhibitor of the NF-kappa B transcription factor. *Science* **242,** 540-546.

Baltimore, D. (1970). Viral RNA-dependent DNA polymerase? *Nature* **226,** 1209-1211.

Barillari, G., Gendelman, R., Gallo, R.C. & Ensoli, B. (1993). The Tat protein of human immunodeficiency virus type 1, a growth factor for AIDS Kaposi sarcoma and cytokine-activated vascular cells, induces adhesion of the same cell types by using integrin receptors recognizing the RGD amino acid sequence. *Proc. Natl. Acad. Sci. USA.* **90,** 7941-7945.

Barin, F., M'boup, S., Denis,F., Kanki, P., Allan, J.S., Lee, T.H. & Essex, M. (1985). Serological evidence for virus related to simian T-lymphotropic retrovirus III in residents of West Africa. *Lancet* **2,** 1387-1389.

Barnett, R.W., Delling, U., Kuperman, R., Sonenberg, N. & Sumner-Smith, M. (1993). Rotational symmetry in ribonucleotide strand requirements for binding of HIV-1 Tat protein to TAR RNA. *Nucleic Acids Res.* **21,** 151-154.

Barre-sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Brun-Vezinet, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220,** 868-871.

Bassuk, A.G., Anandappa, R.T. & Leiden, J.M. (1997). Physical interactions between Ets and NF-kappaB/NFAT proteins play an important role in their cooperative activation of the human immunodeficiency virus enhancer in T cells. *J. Virol.* **71,** 3563-3573.

Beg, A.A. & Baldwin, A.S. Jr. (1993). The I kappa B proteins: multifunctional regulators of Rel/NF-kappa B transcription factors. *Genes Dev.* **7,** 2064-2070.

Ben-David, Y., Giddens, E.B., Letwin, K. & Bernstein, A. (1991). Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes Dev.* **5,** 908-918.

Benkirane, M., Chun, R.F., Xiao, H., Ogryzko, V.V., Howard, B.H., Nakatani, Y. & Jeang, K.T. (1998). Activation of integrated provirus requires histone acetyltransferase. p300 and P/CAF are coactivators for HIV-1 Tat. *J. Biol. Chem.* **273,** 24898-24905.

Berger, S.L. (2002). Histone modifications in transcriptional regulation. *Curr. Opin. Genet. Dev.* **12,** 142-148.

Berk, A.J. (1999). Activation of RNA polymerase II transcription. *Curr. Opin. Cell. Biol.* **11,** 330-335.

Berkhout, B., Gatignol, A., Silver, J. & Jeang, K.T. (1990). Efficient trans-activation by the HIV-2 Tat protein requires a duplicated TAR RNA structure. *Nucleic Acids Res.* **18,** 1839-1846.

Berkhout, B. & Jeang, K.T. (1992). Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* **66,** 139-149.

Berkhout, B., Verhoef, K., van Wamel, J.L. & Back, N.K. (1999). Genetic instability of live, attenuated human immunodeficiency virus type 1 vaccine strains. *J. Virol.* 73, 1138-1145.

Berry, N., Ariyoshi, K., Jobe, O., Ngum, P.T., Corrah, T., Wilkins, A., Whittle, H. & Tedder, R. (1994). HIV type 2 proviral load measured by quantitative polymerase chain reaction correlates with CD4+ lymphopenia in HIV type 2-infected individuals. *AIDS Res. Hum. Retroviruses* **10,** 1031-1037.

Berry, N., Ariyoshi, K., Jaffar, S., Sabally, S., Corrah, T., Tedder, R. & Whittle, H. (1998). Low peripheral blood viral HIV-2 RNA in individuals with high CD4 percentage differentiates HIV-2 from HIV-1 infection. *J. Hum. Virol.* **1,** 457-68.

Berry, N., Ariyoshi, K., Balfe, P., Tedder, R. & Whittle, H. (2001). Sequence specificity of the human immunodeficiency virus type 2 (HIV-2) long terminal repeat U3 region in vivo allows subtyping of the principal HIV-2 viral subtypes A and B. *AIDS Res. Hum. Retroviruses* **17,** 263-267.

Berry, N., Jaffar, S., Van Der Loeff, M.S., Ariyoshi, K., Harding, E., N'Gom, P.T., Dias, F., Wilkins, A., Ricard, D., Aaby, P., Tedder, R. & Whittle, H. (2002). Low level viremia and high CD4% predict normal survival in a cohort of HIV type-2-infected villagers. *AIDS Res. Hum. Retroviruses* **18,** 1167-1173.

Besnier, J.M., Barin, F., Baillou, A., Liard, F., Choutet, P. & Goudeau, A. (1990). Symptomatic HIV-2 primary infection. *Lancet* **335,** 798.

Bharucha, D.C., Zhou, M., Nekhai, S., Brady, J.N., Shukla, R.R. & Kumar, A. (2002). A protein phosphatase from human T cells augments tat transactivation of the human immunodeficiency virus type 1 long-terminal repeat. *Virology* **296,** 6-16.

Bhat, N.K., Komschlies, K.L., Fujiwara, S., Fisher, R.J., Mathieson, B.J., Gregorio, T.A., Young, H.A., Kasik, J.W., Ozato, K. & Papas, T.S. (1989). Expression of ets genes in mouse thymocyte subsets and T cells. *J. Immunol.* **142,** 672-678.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P. & Cullen, B.R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. *EMBO J.* **17,** 7056-7065.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P. & Cullen, B.R. (1999a). Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. *Proc. Natl. Acad. Sci. USA.* **96,** 7791-7796.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P. & Cullen, B.R. (1999b). Highly divergent lentiviral Tat proteins activate viral gene expression by a common mechanism. *Mol. Cell. Biol.* **19,** 4592-4599.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P. & Cullen, B.R. (1999c). Analysis of the effect of natural sequence variation in Tat and in cyclin T on the formation and RNA binding properties of Tat-cyclin T complexes. *J. Virol.* **73,** 5777-5786.

Binley, J.M., Jin, X., Huang, Y., Zhang, L., Cao, Y., Ho, D.D. & Moore, J.P. (1998). Persistent antibody responses but declining cytotoxic T-lymphocyte responses to multiple human immunodeficiency virus type 1 antigens in a long-term nonprogressing individual with a defective p17 proviral sequence and no detectable viral RNA expression. *J. Virol.* **72,** 3472-3474.

Bjorling, E., Scarlatti, G., von Gegerfelt, A., Albert, J., Biberfeld, G., Chiodi, F., Norrby, E. & Fenyo, E.M. (1993). Autologous neutralizing antibodies prevail in HIV-2 but not in HIV-1 infection. *Virology* **193,** 528-530.

Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J. & Bentley, D. (1996). Three functional classes of transcriptional activation domain. *Mol. Cell. Biol.* **16,** 2044-2055.

Bleul, C.C., Farzan, M., Choe, H., Parolin, C., Clark-Lewis, I., Sodroski, J. & Springer, T.A. (1996). The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* **382,** 829-833.

Bourgeois, C.F., Kim, Y.K., Churcher, M.J., West, M.J. & Karn, J. (2002). Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences. *Mol. Cell. Biol.* **22,** 1079-1093.

Brady, M.E., Ozanne, D.M., Gaughan, L., Waite, I., Cook, S., Neal, D.E. & Robson, C.N. (1999). Tip60 is a nuclear hormone receptor coactivator. J. Biol. Chem. **274,** 17599-17604.

Brasseur, R., Cornet, B., Burny, A., Vandenbranden, M. & Ruysschaert, J.M. (1988). Mode of insertion into a lipid membrane of the N-terminal HIV gp41 peptide segment. *AIDS Res. Hum. Retroviruses* **4,** 83-90.

Broder, C.C., Dimitrov, D.S., Blumenthal, R. & Berger, E.A. (1993). The block to HIV-1 envelope glycoprotein-mediated membrane fusion in animal cells expressing human CD4 can be overcome by a human cell component(s). *Virology* **193,** 483-491.

Brodsky, A.S., Erlacher, H.A. & Williamson, J.R. (1998). NMR evidence for a base triple in the HIV-2 TAR C-G.C+ mutant-argininamide complex. *Nucleic Acids Res.* **26,** 1991-1995.

Brown, E.W., Yuhki, N., Packer, C. & O'Brien, S.J. (1994). A lion lentivirus related to feline immunodeficiency virus: epidemiologic and phylogenetic aspects. *J. Virol.* **68,** 5953-5968.

Brownell, J.E. & Allis, C.D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr. Opin. Genet. Dev.* **6,** 176-184.

Brucker, G., Brun-Vezinet, F., Rosenheim, M., Rey, M.A., Katlama, C. & Gentilini, M. (1987). HIV-2 infection in two homosexual men in France. *Lancet* **1,** 223.

Brun-Vezinet, F., Rey, M.A., Katlama, C., Girard, P.M., Roulot, D., Yeni, P., Lenoble, L., Clavel, F., Alizon, M., Gadelle, S., et al. (1987). Lymphadenopathy-associated virus type 2 in AIDS and AIDS-related complex. Clinical and virological features in four patients. *Lancet* **1,** 128-132.

Bruskin, A., Jackson, J., Bishop, J.M., McCarley, D.J. & Schatzman, R.C. (1990). Six amino acids from the retroviral gene gag greatly enhance the transforming potential of the oncogene v-erb-B. *Oncogene* **5,** 15-24.

Bryant, M. & Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc. Natl. Acad. Sci. USA.* **87,** 523-527.

Bryant, M.L., Ratner, L., Duronio, R.J., Kishore, N.S., Devadas, B., Adams, S.P. & Gordon, J.I. (1991). Incorporation of 12-methoxydodecanoate into the human immunodeficiency virus 1 gag polyprotein precursor inhibits its proteolytic processing and virus production in a chronically infected human lymphoid cell line. *Proc. Natl. Acad. Sci. USA.* **88,** 2055-2059.

Buratowski, S. (2000). Snapshots of RNA polymerase II transcription initiation. *Curr. Opin. Cell. Biol.* **12,** 320-325.

Burley, S.K. & Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.* **65,** 769-799.

Butler, J.E. & Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* **16,** 2583-2592.

Bushman, F.D., Fujiwara, T. & Craigie, R. (1990). Retroviral DNA integration directed by HIV integration protein in vitro. *Science* **249,** 1555-1558.

Canonne-Hergaux, F., Aunis, D. & Schaeffer, E. (1995). Interactions of the transcription factor AP-1 with the long terminal repeat of different human immunodeficiency virus type 1 strains in Jurkat, glial, and neuronal cells. *J. Virol.* **69,** 6634-6642.

Carpenter, S., Miller, L.D., Alexandersen, S., Whetstone, C.A., VanDerMaaten, M.J., Viuff, B., Wannemuehler, Y., Miller, J.M. & Roth, J.A. (1992). Characterization of early pathogenic effects after experimental infection of calves with bovine immunodeficiency-like virus. *J. Virol.* **66,** 1074-1083.

Carrington, M., Nelson, G.W., Martin, M.P., Kissner, T., Vlahov, D., Goedert, J.J., Kaslow, R., Buchbinder, S., Hoots, K. & O'Brien, S.J. (1999). HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283,** 1748-1752.

Carroll, R., Peterlin, B.M. & Derse, D. (1992). Inhibition of human immunodeficiency virus type 1 Tat activity by coexpression of heterologous trans activators. *J. Virol.* **66,** 2000-2007.

Carvalho, M., Kirkland, M. & Derse, D. (1993). Protein interactions with DNA elements in variant equine infectious anemia virus enhancers and their impact on transcriptional activity. *J. Virol.* **67,** 6586-6595.

Catucci, M., Venturi, G., Romano, L., Valensin, P.E. & Zazzi, M. (2000). Analysis of the HIV-1 nef gene in five intravenous drug users with long-term nonprogressive HIV-1 infection in Italy. *J. Med. Virol.* **60,** 294-299.

Cavaco-Silva, P., Taveira, N.C., Rosado, L., Lourenco, M.H., Moniz-Pereira, J., Douglas, N.W., Daniels, R.S. & Santos-Ferreira, M.O. (1998). Virological and molecular demonstration of human immunodeficiency virus type 2 vertical transmission. *J. Virol.* **72,** 3418-3422.

Chakrabarti, L., Guyader, M., Alizon, M., Daniel, M.D., Desrosiers, R.C., Tiollais, P. & Sonigo, P. (1987). Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. *Nature* **328,** 543-547.

Champalimaud, J. & Montagnier, L. (1986a). Isolation of a new human retrovirus from West African patients with AIDS. *Science* **223,** 343-346.

Chan, D.C., Fass, D., Berger, J.M. & Kim, P.S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89,** 263-273.

Chan, H.M. & La Thangue, N.B. (2001). p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.* **114,** 2363-2373.

Chang, D.D. & Sharp, P.A. (1989). Regulation by HIV Rev depends upon recognition of splice sites. *Cell* **59,** 789-795.

Chang, Y.N. & Jeang, K.T. (1992). The basic RNA-binding domain of HIV-2 Tat contributes to preferential trans-activation of a TAR2-containing LTR. *Nucleic Acids Res.* **20,** 5465-5472.

Chang, H.K., Gallo, R.C. & Ensoli, B. (1995). Regulation of Cellular Gene Expression and Function by the Human Immunodeficiency Virus Type 1 Tat Protein. *J. Biomed. Sci.* **2,** 189-202.

Chang, L.J., McNulty, E. & Martin, M. (1993). Human immunodeficiency viruses containing heterologous enhancer/promoters are replication competent and exhibit different lymphocyte tropisms. *J. Virol.* **67,** 743-752.

Charneau, P., Borman, A.M., Quillent, C., Guetard, D., Chamaret, S., Cohen, J., Remy, G., Montagnier, L. & Clavel, F. (1994). Isolation and envelope sequence of a highly divergent HIV-1 isolate: definition of a new HIV-1 group. *Virology* **205,** 247-253.

Chen, Z., Telfier, P., Gettie, A., Reed, P., Zhang, L., Ho, D.D. & Marx, P.A. (1996). Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. *J. Virol.* **70,** 3617-3627.

Chen, Z., Luckay, A., Sodora, D.L., Telfer, P., Reed, P., Gettie, A., Kanu, J.M., Sadek, R.F., Yee, J., Ho, D.D., Zhang, L. & Marx, P.A. (1997). Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J. Virol.* **71,** 3953-3960.

Chen, Z., Zhou, P., Ho, D.D., Landau, N.R. & Marx, P.A. (1997). Genetically divergent strains of simian immunodeficiency virus use CCR5 as a coreceptor for entry. *J. Virol.***71,** 2705-2714.

Chen, D., Fong, Y. & Zhou, Q. (1999). Specific interaction of Tat with the human but not rodent P-TEFb complex mediates the species-specific Tat activation of HIV-1 transcription. *Proc. Natl. Acad. Sci. USA.* **96,** 2728-2733.

Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P.D., Wu, L., Mackay, C.R., LaRosa, G., Newman, W., Gerard, N., Gerard, C. & Sodroski, J. (1996). The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* **85,** 1135-1148.

Chun, R.F. & Jeang, K.T. (1996). Requirements for RNA polymerase II carboxyl-terminal domain for activated transcription of human retroviruses human T-cell lymphotropic virus I and HIV-1. J. Biol. Chem. **271,** 27888-27894.

Churcher, M.J., Lamont, C., Hamy, F., Dingwall, C., Green, S.M., Lowe, A.D., Butler, J.G., Gait, M.J. & Karn, J. (1993). High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.* **230,** 90-110.

Clapham, P.R., McKnight, A. & Weiss, R.A. (1992). Human immunodeficiency virus type 2 infection and fusion of CD4-negative human cell lines: induction and enhancement by soluble CD4. *J. Virol.* **66,** 3531-3537.

Clark, N.M., Smith, M.J., Hilfinger, J.M. & Markovitz, D.M. (1993). Activation of the human T-cell leukemia virus type I enhancer is mediated by binding sites for Elf-1 and the pets factor. *J. Virol.* **67,** 5522-5528.

Clark, N.M., Hannibal, M.C. & Markovitz, D.M. (1995). The peri-κB site mediates human immunodeficiency virus type 2 enhancer activation in monocytes but not in T cells. *J. Virol.* **69,** 4854-4861.

Clark, S.J., Saag, M.S., Decker, W.D., Campbell-Hill, S., Roberson, J.L., Veldkamp, P.J., Kappes, J.C., Hahn, B.H. & Shaw, G.M. (1991). High titers of cytopathic virus in plasma of patients with symptomatic primary HIV-1 infection. *N. Engl. J. Med.* **324,** 954-960.

Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M.A., Santos-Ferreira M.O., Laurent, A.G., Dauguet, C., Katlama, C., Rouzioux, C., Klatzmann, D., Champalimaud, J. & Montagnier, L. (1986a). Isolation of a new human retrovirus from West African patients with AIDS. *Science* **223,** 343-346.

Clavel, F., Guyader, M., Guetard, D., Salle, M., Montagnier, L. & Alizon, M. (1986b). Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature* **324,** 691-695.

Clavel, F., Mansinho, K., Chamaret, S., Guetard, D., Favier, V., Nina, J., Santos-Ferreira, M.O., Champalimaud, J.L. & Montagnier, L. (1987a). Human immunodeficiency virus type 2 infection associated with AIDS in West Africa. *N. Engl. J. Med.* **316,** 1180-1185.

Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P. & Weiss, R. (1986). Human immunodeficiency viruses. *Science* **232,** 697.

Coffin, J.M. (1992). Structure and Classification of Retroviruses. *The Retroviridae.* New York: Plenum Press, 19-50.

Coffin, J.M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* **267,** 483-489.

Cohen, E.A., Subbramanian, R.A. & Gottlinger, H.G. (1996). Role of auxiliary proteins in retroviral morphogenesis. *Curr. Top. Microbiol. Immunol.* **214,** 219-235.

Coleman, R.A., Taggart, A.K., Burma, S., Chicca, J.J. 2nd & Pugh, B.F. (1999). TFIIA regulates TBP and TFIID dimers. *Mol. Cell.* **4,** 451-457.

Connor, R.I. & Ho, D.D. (1994). Human immunodeficiency virus type 1 variants with increased replicative capacity develop during the asymptomatic stage before disease progression. *J. Virol.* **68,** 4400-4408.

Control, C.f.D. (1981a). Karposi's sarcoma and *Pneumocystis* pneumonia among homosexual men – New York City and California. *MMWR* **30.**

Control, C.f.D. (1981b). *Pneumocystis* pneumonia-Los Angeles. *MMWR* **30,** 250-252.

Cooper, D.A., Gold, J., Maclean, P., Donovan, B., Finlayson, R., Barnes, T.G., Michelmore, H.M., Brooke, P. & Penny, R. (1985). Acute AIDS retrovirus infection: definition of a clinical illness associated with seroconversion. *Lancet* **8428,** 537-540.

Corboy, J.R., Buzy, J.M., Zink, M.C. & Clements, J.E. (1992). Expression directed from HIV long terminal repeats in the central nervous system of transgenic mice. *Science* **258,** 1804-1808.

Corboy, J.R. & Garl, P.J. (1997). HIV-1 LTR DNA sequence variation in brain-derived isolates. *J. Neurovirol.* **3,** 331-341.

Cortes, E., Detels, R., Aboulafia, D., Li, X.L., Moudgil, T., Alam, M., Bonecker, C., Gonzaga, A., Oyafuso, L. & Tondo, M. (1989). HIV-1, HIV-2, and HTLV-1 infection in high-risk groups in Brazil. *N. Engl. J. Med.* **320,** 953-958.

Coulombe, B. & Burton, Z.F. (1999). DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol.* Rev. **63,** 457-478.

Craigie, R., Fujiwara, T. & Bushman, F. (1990). The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. *Cell* **62,** 829-837.

Cujec, T.P., Okamoto, H., Fujinaga, K., Meyer, J., Chamberlin, H., Morgan, D.O. & Peterlin, B.M. (1997). The HIV transactivator TAT binds to the CDK-activating kinase and activates the phosphorylation of the carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11,** 2645-2657.

Cullen, B.R., Lomedico, P.T. & Ju, G. (1984). Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis. *Nature* **307,** 241-245.

Daar, E.S., Moudgil, T., Meyer, R.D. & Ho, D.D. (1991). Transient high levels of viremia in patients with primary human immunodeficiency virus type 1 infection.*N. Engl. J. Med.* **324,** 961-964.

Dalgleish, A.G., Beverley, P.C., Clapham, P.R., Crawford, D.H., Greaves, M.F. & Weiss, R.A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312,** 763-767.

Daniel, M.D., Letvin, N.L., King, N.W., Kannagi, M., Sehgal, P.K., Hunt, R.D., Kanki, P.J., Essex, M. & Desrosiers, R.C. (1985). Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. *Science* **228,** 1201-1204.

Daniel, M.D., Kirchhoff, F., Czajak, S.C., Sehgal, P.K. & Desrosiers, R.C. (1992). Protective effects of a live attenuated SIV vaccine with a deletion in the nef gene. *Science* **258,** 1938-1941.

Das, A.T., Klaver, B. & Berkhout, B. (1998). The 5' and 3' TAR elements of human immunodeficiency virus exert effects at several points in the virus life cycle. *J. Virol.* **72,** 9217-9223.

Das, A.T., Klaver, B. & Berkhout, B. (1999). A hairpin structure in the R region of the human immunodeficiency virus type 1 RNA genome is instrumental in polyadenylation site selection. *J. Virol.* **73,** 81-91.

Davis, B.D., Dulbecco, R., Eisen, H.N. & Ginsberg, H.S. (1990). Microbiology, Fourth Edition: J.B. Lippincott Company.

Deacon, N.J., Tsykin, A., Solomon, A., Smith, K., Ludford-Menting, M., Hooker, D.J., McPhee, D.A., Greenway, A.L., Ellett, A., Chatfield, C., et al. (1995). Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science* **270,** 988-991.

Dean, M., Carrington, M., Winkler, C., Huttley, G.A., Smith, M.W., Allikmets, R., Goedert, J.J., Buchbinder, S.P., Vittinghoff, E., Gomperts, E., Donfield, S., Vlahov, D., Kaslow, R., Saah, A., Rinaldo, C., Detels, R. & O'Brien, S.J. (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* **273,** 1856-1862.

De Cock, K.M. & Brun Vezinet, F. (1989). Epidemiology of HIV-2 infection. *AIDS* **3,** Suppl 1:S89-95.

De Cock, K.M., Odehouri, K., Colebunders, R.L., Adjorlolo, G., Lafontaine, M.F., Porter, A., Gnaore, E., Diaby, L., Moreau, J., Heyward, W.L., et al. (1990). A comparison of HIV-1 and HIV-2 infections in hospitalized patients in Abidjan, Cote d'Ivoire. *AIDS* **4,** 443-448.

Delaporte, E., Janssens, W., Peeters, M., Buve, A., Dibanga, G., Perret, J.L., Ditsambou, V., Mba, J.R., Courbot, M.C., Georges, A., Bourgeois, A., Samb, B., Henzel, D., Heyndrickx, L., Fransen, K., van der Groen, G. & Larouze, B. (1996). Epidemiological and molecular characteristics of HIV infection in Gabon, 1986-1994. *AIDS* **10,** 903-910.

De Leys, R., Vanderborght, B., Vanden Haesevelde, M., Heyndrickx, L., Van Geel, A., Wauters, C., Bernaerts, R., Saman, E., Nijs, P., Willems, B., et al. (1990). Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. *J. Virol.* **64,** 1207-1216.

Delassus, S., Cheynier, R., & Wain-Hobson, S. (1991). Evolution of human immunodeficiency virus type 1 *nef* and long terminal repeat sequences over 4 years *in vivo* and *in vitro*. *J. Virol.* **65,** 225-231.

Delassus, S., Meyerhans, A., Cheynier, R. & Wain-Hobson, S. (1992). Absence of selection of HIV-1 variants in vivo based on transcription/transactivation during progression to AIDS. *Virology* **188,** 811-818.

Del Mistro, A., Chotard, J., Hall, A.J., Whittle, H., De Rossi, A. & Chieco-Bianchi, L. (1992). HIV-1 and HIV-2 seroprevalence rates in mother-child pairs living in The Gambia (west Africa). *J.Acquir.Immune.Defic.Syndr.* **5,** 19-24.

Delwart, E.L., Sheppard, H.W., Walker, B.D., Goudsmit, J. & Mullins, J.I. (1994). Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. *J. Virol.* **68,** 6672-6683.

Deng, L., Wang, D., de la Fuente, C., Wang, L., Li, H., Lee, C.G., Donnelly, R., Wade, J.D., Lambert, P. & Kashanchi, F. (2001). Enhancement of the p300 HAT activity by HIV-1 Tat on chromatin DNA. *Virology* **289,** 312-326.

Desrosiers, R.C. (1990). The simian immunodeficiency viruses. *Annu. Rev. Immunol.* **8,** 557-578.

Desrosiers, R.C. & Letvin, N.L. (1987). Animal models for aquired immunodeficiency syndrome. *Reviews of Infectious Diseases* **9,** 438-446.

Desrosiers, R.C. (1999). Strategies used by human immunodeficiency virus that allow persistent viral replication. *Nat. Med.* **5,** 723-725.

DeStefano, J.J., Buiser, R.G., Mallaber, L.M., Bambara, R.A. & Fay, P.J. (1991). Human immunodeficiency virus reverse transcriptase displays a partially processive 3' to 5' endonuclease activity. *J. Biol. Chem.* 266, 24295-24301.

de Wolf, F., Spijkerman, I., Schellekens, P.T., Langendam, M., Kuiken, C., Bakker, M., Roos, M., Coutinho, R., Miedema, F. & Goudsmit, J. (1997). AIDS prognosis based on HIV-1 RNA, CD4+ T-cell count and function: markers with reciprocal predictive value over time after seroconversion. *AIDS* **11,** 1799-1806.

Dietrich, U., Adamski, M., Kreutz, R., Seipp, A., Kuhnel, H. & Rubsamen-Waigmann, H. (1989). A highly divergent HIV-2-related isolate. *Nature* **342,** 948-950.

Dillon, P.J., Nelbock, P., Perkins, A. & Rosen, C.A. (1991). Structural and functional analysis of the human immunodeficiency virus type 2 Rev protein. *J. Virol.* **65,** 445-449.

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M. & Skinner, M.A. (1990). HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *EMBO J.* **9,** 4145-4153.

Dittmer, U., Petry, H., Stahl, H.C., Nisslein, T., Spring, M., Luke, W., Bodemer, W., Kaup, F.J. & Hunsmann, G. (1996). T cell apoptosis in human immunodeficiency virus type 2 and simian immunodeficiency virus-infected macaques. *J. Gen. Virol.* **77,** 2433-2436.

Djomand, G., Greenberg, A.E., Sassan-Morokro, M., Tossou, O., Diallo, M.O., Ekpini, E., Ghys, P., Soro, B., Brattegaard, K., Yapi, A., et al. (1995). The epidemic of HIV/AIDS in Abidjan, Cote d'Ivoire: a review of data collected by Projet RETRO-CI from 1987 to 1993. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **10,** 358-365.

Doerr, H.W., Selb, B., Brede, H.D., Biesert, L., Von Briesen, H. & Rubsamen-Waigmann H. (1987). HIV-2 infection also in Germany. *Dtsch. Med. Wochenschr.* **112,** 1237.

Dorrell, L., Willcox, B.E., Jones, E.Y., Gillespie, G., Njai, H., Sabally, S., Jaye, A., DeGleria, K., Rostron, T., Lepin, E., McMichael, A., Whittle, H. & Rowland-Jones, S. (2001). Cytotoxic T lymphocytes recognize structurally diverse, clade-specific and cross-reactive peptides in human immunodeficiency virus type-1 gag through HLA-B53. *Eur. J. Immunol.* **31,** 1747-1756.

Dragic, T., Charneau, P., Clavel, F. & Alizon, M. (1992). Complementation of murine cells for human immunodeficiency virus envelope/CD4-mediated fusion in human/murine heterokaryons. *J. Virol.* **66,** 4794-4802.

Dufoort, G., Courouce, A.M., Ancelle-Park, R. & Bletry, O. (1988). No clinical signs 14 years after HIV-2 transmission via blood transfusion. *Lancet* **2,** 510.

Eberharter, A. & Becker, P.B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO* **3,** 224-229.

Edinger, A.L., Amedee, A., Miller, K., Doranz, B.J., Endres, M., Sharron, M., Samson, M., Lu, Z.H., Clements, J.E., Murphey-Corb, M., Peiper, S.C., Parmentier, M., Broder, C.C. & Doms, R.W. (1997). Differential utilization of CCR5 by macrophage and T cell tropic simian immunodeficiency virus strains. *Proc. Natl. Acad. Sci. USA.* **94,** 4005-4010.

Elangovan, B., Subramanian, T. & Chinnadurai, G. (1992). Functional comparison of the basic domains of the Tat proteins of human immunodeficiency virus types 1 and 2 in trans activation. *J. Virol.* **66,** 2031-2036.

El Kharroubi, A. & Martin, M.A. (1996). cis-acting sequences located downstream of the human immunodeficiency virus type 1 promoter affect its chromatin structure and transcriptional activity. *Mol. Cell. Biol.* **16,** 2958-2966.

El Kharroubi, A., Piras, G., Zensen, R. & Martin, M.A. (1998). Transcriptional activation of the integrated chromatin-associated human immunodeficiency virus type 1 promoter. *Mol. Cell. Biol.* **18,** 2535-2544.

Emerman, M., Guyader, M., Montagnier, L., Baltimore, D. & Muesing, M.A. (1987). The specificity of the human immunodeficiency virus type 2 transactivator is different from that of human immunodeficiency virus type 1. *EMBO J.* **6,** 3755-3760.

Endres, M.J., Clapham, P.R., Marsh, M., Ahuja, M., Turner, J.D., McKnight, A., Thomas, J.F., Stoebenau-Haggarty, B., Choe, S., Vance, P.J., Wells, T.N., Power, C.A., Sutterwala, S.S., Doms, R.W., Landau, N.R. & Hoxie, J.A. (1996). CD4-independent infection by HIV-2 is mediated by fusin/CXCR4. *Cell* **87,** 745-756.

Engelman, A., Mizuuchi, K. & Craigie, R. (1991). HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* **67,** 1211-1221.

Englund, G., Hoggan, M.D., Theodore, T.S. & Martin, M.A. (1991). A novel HIV-1 isolate containing alterations affecting the NF-kappa B element. *Virology* **181,** 150-157.

Ensoli, B., Buonaguro, L., Barillari, G., Fiorelli, V., Gendelman, R., Morgan, R.A., Wingfield, P. & Gallo, R.C. (1993). Release, uptake, and effects of extracellular human immunodeficiency virus type 1 Tat protein on cell growth and viral transactivation. *J. Virol.* **67,** 277-287.

Estable, M.C., Bell, B., Merzouki, A., Montaner, J.S., O'Shaughnessy, M.V., & Sadowski, I.,J. (1996). Human immunodeficiency virus type 1 long terminal repeat variants from 42 patients representing all stages of infection display a wide range of sequence polymorphism and transcriptional activity. *J. Virol.* **70,** 4053-4062.

Evans, B.G., Gill, O.N., Gleave, S.R., Mortimer, P.P. & Parry, J.V. (1991). HIV-2 in the United Kingdom--a review. *CDR (Lond Engl Rev)* **1,** R19-23.

Fang, G., Burger, H., Chappey, C., Rowland-Jones, S., Visosky, A., Chen, C.H., Moran, T., Townsend, L., Murray, M. & Weiser, B. (2001). Analysis of transition from long-term nonprogressive to progressive infection identifies sequences that may attenuate HIV type 1. *AIDS Res. Hum. Retroviruses* **17,** 1395-1404.

Farr, A. & Roman, A. (1992). A pitfall of using a second plasmid to determine transfection efficiency. *Nucleic Acids Res.* **20,** 920.

Fass, D., Davey, R.A., Hamson, C.A., Kim, P.S., Cunningham, J.M. & Berger, J.M. (1997). Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution. *Science* **277,** 1662-1666.

Faulkner, N.E., Hilfinger, J.M. & Markovitz, D.M. (2001). Protein phosphatase 2A activates the HIV-2 promoter through enhancer elements that include the pets site. *J. Biol. Chem.* **276,** 25804-25812.

Faye, A., Burgard, M., Crosnier, H., Retbi, J.M. & Blanche, S. (1997). Human immunodeficiency virus type 2 infection in children. *J. Pediatr.* **130,** 994-997.

Feinberg, M.B., Baltimore, D. & Frankel, A.D. (1991). The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proc. Natl. Acad. Sci. USA.* **88,** 4045-4049.

Feng, Y., Broder, C.C., Kennedy, P.E. & Berger, E.A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **272,** 872-877.

Fenrick, R., Malim, M.H., Hauber, J., Le, S.Y., Maizel, J. & Cullen, B.R. (1989). Functional analysis of the Tat trans activator of human immunodeficiency virus type 2. *J. Virol.* **63,** 5006-5012.

Finzi, D. & Siliciano, R.F. (1998). Viral dynamics in HIV-1 infection. *Cell* **93,** 665-671.

Franchini, G., Gurgo, C., Guo, H.G., Gallo, R.C., Collalti, E., Fargnoli, K.A., Hall, L.F., Wong-Staal, F. & Reitz, M.S. Jr. (1987). Sequence of simian immunodeficiency virus and its relationship to the human immunodeficiency viruses. *Nature* **328,** 539-543.

Frankel, A.D. (1992). Activation of HIV transcription by Tat. *Curr. Opin. Genet. Dev.* **2,** 293-298.

Fu, G.K. & Markovitz, D.M. (1996). Purification of the pets factor. *J. Biol. Chem.* **271,** 19599-19605.

Fu, G.K., Grosveld, G. & Markovitz, D.M. (1997). DEK, an autoantigen involved in a chromosomal translocation in acute myelogenous leukemia, binds to the HIV-2 enhancer. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 1811-1815.

Fujii, S., Obaru, K., Matsushita, S., Morikita, T., Higuchi, H., Fujimoto, K. & Takatsuki, K. (1997). Characterization of proviral DNA from an individual with long-term, nonprogressive infection with HIV-1 and nonrecoverable virus. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **15,** 247-256.

Fujinaga, K., Cujec, T.P., Peng, J., Garriga, J., Price, D.H., Grana, X. & Peterlin, B.M. (1998). The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. *J. Virol.* **72,** 7154-7159.

Fujinaga, K., Taube, R., Wimmer, J., Cujec, T.P. & Peterlin, B.M. (1999). Interactions between human cyclin T, Tat, and the transactivation response element (TAR) are disrupted by a cysteine to tyrosine substitution found in mouse cyclin T. *Proc. Natl. Acad. Sci. USA.* **96,** 1285-1290.

Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M. & Nishida, E. (1997). CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390,** 308-311.

Fung, Y.K., Fadly, A.M., Crittenden, L.B. & Kung, H.J. (1981). On the mechanism of retrovirus-induced avian lymphoid leukosis: deletion and integration of the proviruses. *Proc. Natl. Acad. Sci. USA.* **78,** 3418-3422.

Gaines, H., Albert, J., Von Sydow, M., Sonnerborg, A., Chiodi, F., Ehrnst, A., Strannegard, O. & Asjo, B. (1987). HIV antigenaemia and virus isolation from plasma during primary HIV infection. *Lancet* **1,** 1317-1318.

Gait, M.J. & Karn, J. (1993). RNA recognition by the human immunodeficiency virus Tat and Rev proteins. *Trends. Biochem. Sci.* **18,** 255-259.

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G,M., Kaplan, M., Haynes, B.F., Palker, R., Redfield, R., Oleske, J. & Safai, B. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk os AIDS. *Science* **224,** 500-503.

Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., Leibowitch, J. & Popovic, M. (1983). Isolation of human T-cell leukaemia virus in acquired immune deficiency syndrome (AIDS). *Science* **220,** 865-867.

Gao, F., Yue, L., White, A.T., Pappas, P.G., Barchue, J., Hanson, A.P., Greene, B.M., Sharp, P.M., Shaw, G.M. & Hahn, B.H. (1992). Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. *Nature* **358,** 495-499.

Gao, F., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., et al. (1994). Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *J. Virol.* **68,** 7433-7447.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., Sharp, P.M. & Hahn, B.H. (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **397,** 436-441.

Garber, M.E., Wei, P. & Jones, K.A. (1998a). HIV-1 Tat interacts with cyclin T1 to direct the P-TEFb CTD kinase complex to TAR RNA. *Cold Spring Harb. Symp. Quant. Biol.* **63,** 371-380.

Garber, M.E., Wei, P., KewalRamani, V.N., Mayall, T.P., Herrmann, C.H., Rice, A.P., Littman, D.R. & Jones, K.A. (1998b). The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes Dev.* **12,** 3512-3527.

Garcia, J.A., Harrich, D., Soultanakis, E., Wu, F., Mitsuyasu, R. & Gaynor, R.B. (1989). Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. *EMBO J.* **8,** 765-778.

Garcia-Martinez, L.F., Mavankal, G., Peters, P., Wu-Baer, F. & Gaynor, R.B. (1995). Tat functions to stimulate the elongation properties of transcription complexes paused by the duplicated TAR RNA element of human immunodeficiency virus 2. *J. Mol. Biol.* **254,** 350-363.

Garcia-Martinez, L.F., Ivanov, D. & Gaynor, R.B. (1997a). Association of Tat with purified HIV-1 and HIV-2 transcription preinitiation complexes. *J. Biol. Chem.* **272,** 6951-6958.

Garcia-Martinez, L.F., Mavankal, G., Neveu, J.M., Lane, W.S., Ivanov, D. & Gaynor, R.B. (1997b). Purification of a Tat-associated kinase reveals a TFIIH complex that modulates HIV-1 transcription. *EMBO J.* **16,** 2836-2850.

Gayle, H.D., Gnaore, E., Adjorlolo, G., Ekpini, E., Coulibaly, R., Porter, A., Braun, M.M., Zabban, M.L., Andou, J., Timite, A., et al. (1992). HIV-1 and HIV-2 infection in children in Abidjan, Cote d'Ivoire. *J.Acquir.Immune.Defic.Syndr.* **5,** 513-517.

Gaynor, R. (1992). Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* **6,** 347-363.

Geffin, R., Wolf, D., Muller, R., Hill, M.D., Stellwag, E., Freitag, M., Sass, G., Scott, G.B. & Baur, A.S. (2000). Functional and structural defects in HIV type 1 nef genes derived from pediatric long-term survivors. *AIDS Res. Hum. Retroviruses* **16,** 1855-1868.

Gelderblom, H., Bauer, H., Bolognesi, D.P. & Frank, H. (1972). Morphogenesis and structure of RNA tumor viruses. Electron microscopy studies on C-type viruses. *Zentralbl Bakteriol [Orig A]* **220,** 79-90.

Gendelman, H.E., Orenstein, J.M., Baca, L.M., Weiser, B., Burger, H., Kalter, D.C. & Meltzer, M.S. (1989). The macrophage in the persistence and pathogenesis of HIV infection. *AIDS* **3,** 475-495.

Geyer, M., Fackler, O.T. & Peterlin, B.M. (2001). Structure--function relationships in HIV-1 Nef. *EMBO Rep.* **2,** 580-585.

Ghys, P.D., Fransen, K., Diallo, M.O., Ettiegne-Traore, V., Coulibaly, I.M., Yeboue, K.M., Kalish, M.L., Maurice, C., Whitaker, J.P., Greenberg, A.E. & Laga, M. (1997). The associations between cervicovaginal HIV shedding, sexually transmitted diseases and immunosuppression in female sex workers in Abidjan, Cote d'Ivoire. *AIDS* **11,** F85-93.

Gill, G. (2001). Regulation of the initiation of eukaryotic transcription. *Essays Biochem.* **37,** 33-43.

Goff, S.P. (1990). Retroviral reverse transcriptase: synthesis, structure, and function. *J. Acquir. Immune Defic. Syndr.* **3,** 817-831.

Goff, S.P. (1992). Genetics of retroviral integration. *Annu. Rev. Genet.* **26,** 527-544.

Gojobori, T., Moriyama, E.N., Ina, Y., Ikeo, K., Miura, T., Tsujimoto, H., Hayami, M. & Yokoyama, S. (1990). Evolutionary origin of human and simian immunodeficiency viruses. *Proc. Natl. Acad. Sci. USA.* **87,** 4108-4111.

Golub, E.I., Li, G.G. & Volsky, D.J. (1990). Differences in the basal activity of the long terminal repeat determine different replicative capacities of two closely related human immunodeficiency virus type 1 isolates. *J. Virol.* **64,** 3654-3660.

Gomez-Roman, V.R., Vazquez, J.A., del Carmen Basualdo, M., Estrada, F.J., Ramos-Kuri, M., & Soler, C. (2000). nef/long terminal repeat quasispecies from HIV type 1-infected Mexican patients with different progression patterns and their pathogenesis in hu-PBL-SCID mice. *AIDS Res. Hum. Retroviruses* **16,** 441-452.

Goodenow, M.M. & Hayward, W.S. (1987). 5' long terminal repeats of myc-associated proviruses appear structurally intact but are functionally impaired in tumors induced by avian leukosis viruses. *J. Virol.* **61,** 2489-2498.

Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J. & Wain-Hobson, S. (1989). HIV-1 isolates are rapidly evolving quasispecies: evidence for viral mixtures and preferred nucleotide substitutions. *J. Acquir. Immune Defic. Syndr.* **2,** 344-352.

Goodrich, J.A. & Tjian, R. (1994). TBP-TAF complexes: selectivity factors for eukaryotic transcription. *Curr. Opin. Cell Biol.* **6,** 403-409.

Gorelick, R.J., Henderson, L.E., Hanser, J.P. & Rein, A. (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a "zinc finger-like" protein sequence. *Proc. Natl. Acad. Sci. USA.* **85,** 8420-8424.

Gorelick, R.J., Nigida, S.M. Jr., Bess, J.W. Jr., Arthur, L.O., Henderson, L.E. & Rein, A. (1990). Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* **64,** 3207-3211.

Gorrell, M.D., Brandon, M.R., Sheffer, D., Adams, R.J. & Narayan, O. (1992). Ovine lentivirus is macrophagetropic and does not replicate productively in T lymphocytes. *J. Virol.* **66,** 2679-2688.

Gotch, F., McAdam, S.N., Allsopp, C.E., Gallimore, A., Elvin, J., Kieny, M.P., Hill, A.V., McMichael, A.J. & Whittle, H.C. (1993). Cytotoxic T cells in HIV2 seropositive Gambians. Identification of a virus-specific MHC-restricted peptide epitope. *J. Immunol.* **151,** 3361-3369.

Gottlieb, G.J. & Ackerman, A.B. (1982). Kaposi's sarcoma: an extensively disseminated form in young homosexual men. *Hum. Pathol.* **13,** 882-892.

Gottlinger, H.G., Sodroski, J.G. & Haseltine, W.A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA.* **86,** 5781-5785.

Grankvist, O., Bredberg-Raden, U., Gustafsson, A., Albert, J., Albino, P., Andreasson, P.A., Naucler, A., Biberfeld, G. & Wadell, G. (1992). Improved detection of HIV-2 DNA in clinical samples using a nested primer-based polymerase chain reaction. *J. Acquir. Immune Defic. Syndr.* **5,** 286-293.

Grant, A.D., Djomand, G. & De Cock, K.M. (1997). Natural history and spectrum of disease in adults with HIV/AIDS in Africa. *AIDS* **11,** Suppl. B S43-54.

Grassly, N.C., Xiang, Z., Ariyoshi, K., Aaby, P., Jensen, H., van der Loeff, M.S., Dias, F., Whittle, H. & Breuer, J. (1998). Mortality among human immunodeficiency virus type 2-positive villagers in rural Guinea-Bissau is correlated with viral genotype. *J. Virol.* **72,** 7895-7899.

Greenberg, A., Coulibaly, I.M., Kadio, A, et al. (1997). Trends in the HIV-1 and HIV-2 epidemics in Abidjan, Cote d'Ivoire: 11 years of HIV serosurveillance at Project Retro-CI. *Xth International Conference on AIDS and STD in Africa, Abidjan. [Abstract B.041].*

Greenberg, M., DeTulleo, L., Rapoport, I., Skowronski, J. & Kirchhausen, T. (1998). A dileucine motif in HIV-1 Nef is essential for sorting into clathrin-coated pits and for downregulation of CD4. *Curr. Biol.* **8,** 1239-1242.

Grez, M., Dietrich, U., Balfe, P., Von Briesen, H., Maniar, J.K., Mahambre, G., Delwart, E.L., Mullins, J.I. & Rubsamen-Waigmann, H. (1994). Genetic analysis of human immunodeficiency virus type 1 and

2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses. *J. Virol.* **68,** 2161-2168.

Griffin, G.E., Leung, K., Folks, T.M., Kunkel, S. & Nabel, G.J. (1989). Activation of HIV gene expression during monocyte differentiation by induction of NF-kappa B. *Nature* **339,** 70-73.

Guillon, C., Van der Ende, M.E., Boers, P.H., Gruters, R.A., Schutten, M. & Osterhaus, A.D. (1998). Coreceptor usage of human immunodeficiency virus type 2 primary isolates and biological clones is broad and does not correlate with their syncytium-inducing capacities. *J. Virol.* **72,** 6260-6263.

Guntaka, R.V. (1993). Transcription termination and polyadenylation in retroviruses. *Microbiol. Rev.* **57,** 511-521.

Gurtler, L.G., Hauser, P.H., Eberle, J., Von Brunn, A., Knapp, S., Zekeng, L., Tsague, J.M. & Kaptue, L. (1994). A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. *J. Virol.* **68,** 1581-1585.

Guyader, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L. & Alizon, M. (1987). Genome organisation and transactivation of the human immunodeficiency virus type 2. *Nature* **326,** 662-669.

Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L.U., Green, M. & Reinberg, D. (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. *Genes Dev.* **7,** 1021-1032.

Hahn, B.H., Shaw, G.M., De Cock, K.M. & Sharp, P.M. (2000). AIDS as a zoonosis: scientific and public health implications. *Science* **287,** 607-614.

Hammes, S.R., Dixon, E.P., Malim, M.H., Cullen, B.R. & Greene, W.C. (1989). Nef protein of human immunodeficiency virus type 1: evidence against its role as a transcriptional inhibitor. *Proc. Natl. Acad. Sci. USA.* **86,** 9549-9553.

Hampe, A., Gobet, M., Sherr, C.J. & Galibert, F. (1984). Nucleotide sequence of the feline retroviral oncogene v-fms shows unexpected homology with oncogenes encoding tyrosine-specific protein kinases. *Proc. Natl. Acad. Sci. USA.* **81,** 85-89.

Hannibal, M.C., Markovitz, D.M., Clark, N. & Nabel, G.J. (1993). Differential activation of human immunodeficiency virus type 1 and type 2 transcription by specific T-cell activation signals. *J. Virol.* **67,** 5035-5040.

Hannibal, M.C., Markovitz, D.M. & Nabel, G.J. (1994). Multiple *cis*-acting elements in the human immunodeficiency virus type 2 enhancer mediate the response to T-cell receptor stimulation by antigen in a T-cell hybridoma line. *Blood.* **83,** 1839-1846.

Harrington, R.D. & Geballe, A.P. (1993). Cofactor requirement for human immunodeficiency virus type 1 entry into a CD4-expressing human cell line. *J. Virol.* **67,** 5939-5947.

Harrich, D., Garcia, J., Wu, F., Mitsuyasu, R., Gonazalez, J. & Gaynor, R. (1989). Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* **63,** 2585-2591.

Harrich, D., Garcia, J., Mitsuyasu, R. & Gaynor, R. (1990). TAR independent activation of the human immunodeficiency virus in phorbol ester stimulated T lymphocytes. *EMBO J.* **9,** 4417-4423.

Harrich, D., Ulich, C. & Gaynor, R.B. (1996). A critical role for the TAR element in promoting efficient human immunodeficiency virus type 1 reverse transcription. *J. Virol.* **70,** 4017-4027.

Harris, C., Small, C.B., Klein, R.S., Friedland, G.H., Moll, B., Emeson, E.E., Spigland, I. & Steigbigel, N.H. (1983). Immunodeficiency in female sexual partners of men with the acquired immunodeficiency syndrome. *N. Engl. J. Med.* **308,** 1181-1184.

Hawkes, S., West, B., Wilson, S., Whittle, H. & Mabey, D. (1995). Asymptomatic carriage of Haemophilus ducreyi confirmed by the polymerase chain reaction. *Genitourin. Med.* **71,** 224-227.

Hazan, U., Thomas, D., Alcami, J., Bachelerie, F., Israel, N., Yssel, H., Virelizier, J.L. & Arenzana-Seisdedos, F. (1990). Stimulation of a human T-cell clone with anti-CD3 or tumor necrosis factor induces NF-kappa B translocation but not human immunodeficiency virus 1 enhancer-dependent transcription. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 7861-7865.

He, G., Ylisastigui, L. & Margolis, D.M. (2002). The regulation of HIV-1 gene expression: the emerging role of chromatin. *DNA Cell Biol.* **21,** 697-705.

Hendel, H., Caillat-Zucman, S., Lebuanec, H., Carrington, M., O'Brien, S., Andrieu, J.M., Schachter, F., Zagury, D., Rappaport, J., Winkler, C., Nelson, G.W. & Zagury, J.F. (1999). New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS. *J. Immunol.* **162,** 6942-6946.

Henderson, A.J. & Calame, K.L. (1997). CCAAT/enhancer binding protein (C/EBP) sites are required for HIV-1 replication in primary macrophages but not CD4(+) T cells. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 8714-8719.

Heneine, W., Switzer, W.M., Sandstrom, P., Brown, J., Vedapuri, S., Schable, C.A., Khan, A.S., Lerche, N.W., Schweizer, M., Neumann-Haefelin, D., Chapman, L.E. & Folks, T.M. (1998). Identification of a human population infected with simian foamy viruses. *Nat. Med.* **4,** 403-407.

Heredia, A., Vallejo, A., Soriano, V., Epstein, J.S. & Hewlett, I.K. (1997). Chemokine receptors and HIV-2. *AIDS* **11,** 1198-1199.

Herrmann, C.H. & Rice, A.P. (1993). Specific interaction of the human immunodeficiency virus Tat proteins with a cellular protein kinase. *Virology* **197,** 601-608.

Herrmann, C.H. & Rice, A.P. (1995). Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. *J. Virol.* **69,** 1612-1620.

Hilfinger, J., Clark, N., Smith, M.J., Robinson, K. & Markovitz, D.M. (1993). Differential regulation of the human immunodeficiency virus type 2 enhancer in monocytes at various stages of differentiation. (1993). *J. Virol.* **67,** 4448-4453.

Hill, C.M., Deng, H., Unutmaz, D., Kewalramani, V.N., Bastiani, L., Gorny, M.K., Zolla-Pazner, S. & Littman, D.R. (1997). Envelope glycoproteins from human immunodeficiency virus types 1 and 2 and simian immunodeficiency virus can use human CCR5 as a coreceptor for viral entry and make direct CD4-dependent interactions with this chemokine receptor. *J. Virol.* **71,** 6296-6304.

Hines, R., Sorensen, B.R., Shea, M.A. & Maury, W. (2004). PU.1 binding to ets motifs within the equine infectious anemia virus long terminal repeat (LTR) enhancer: regulation of LTR activity and virus replication in macrophages. *J. Virol.* **78,** 3407-3418.

Hirose, Y. & Manley, J.L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev.* **14,** 1415-1429.

Hirsch, V.M., Dapolito, G., McGann, C., Olmsted, R.A., Purcell, R.H. & Johnson, P.R. (1989a). Molecular cloning of SIV from sooty mangabey monkeys. *J Med Primatol* **18,** 279-285.

Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H. & Johnson, P.R. (1989b). An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* **339,** 389-392.

Ho, D.D., Rota, T.R. & Hirsch, M.S. (1986). Infection of monocyte/macrophages by human T lymphotropic virus type III. *J. Clin. Invest.* **77,** 1712-1715.

Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M. & Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373,** 123-126.

Hoey, T., Weinzierl, R.O., Gill, G., Chen, J.L., Dynlacht, B.D. & Tjian, R. (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* **72,** 247-260.

Horvat, R.T. & Wood, C. (1989). HIV promoter activity in primary antigen-specific human T lymphocytes. *J. Immunol.* **143,** 2745-2751.

Hottiger, M.O. & Nabel, G.J. (1998). Interaction of human immunodeficiency virus type 1 Tat with the transcriptional coactivators p300 and CREB binding protein. *J. Virol.* **72,** 8252-8256.

Howcroft, T.K., Strebel, K., Martin, M.A. & Singer, D.S. (1993). Repression of MHC class I gene promoter activity by two-exon Tat of HIV. *Science* **260,** 1320-1322.

Hsia, S.C. & Shi, Y.B. (2002). Chromatin disruption and histone acetylation in regulation of the human immunodeficiency virus type 1 long terminal repeat by thyroid hormone receptor. *Mol. Cell. Biol.* **22,** 4043-4052.

Hu, W.S. & Temin, H.M. (1990). Retroviral recombination and reverse transcription. *Science* **250,** 1227-1233.

Huang, L.M. & Jeang, K.T. (1993). Increased spacing between Sp1 and TATAA renders human immunodeficiency virus type 1 replication defective: implication for Tat function. *J. Virol.* **67,** 6937-6944.

Huang, L., Bosch, I., Hofmann, W., Sodroski, J. & Pardee, A.B. (1998a). Tat protein induces human immunodeficiency virus type 1 (HIV-1) coreceptors and promotes infection with both macrophage-tropic and T-lymphotropic HIV-1 strains. *J. Virol.* **72,** 8952-8960.

Huang, Y., Zhang, L. & Ho, D,D. (1995a). Characterization of nef sequences in long-term survivors of human immunodeficiency virus type 1 infection. *J. Virol.* **69,** 93-100.

Huang, Y., Zhang, L. & Ho, D.D. (1995b). Biological characterization of *nef* in long-term survivors of human immunodeficiency virus type 1 infection. *J. Virol.* **69,** 8142-8146.

Huang, Y., Zhang, L. & Ho, D.D. (1998b). Characterization of gag and pol sequences from long-term survivors of human immunodeficiency virus type 1 infection. *Virology* **240,** 36-49.

Huet, T., Cheynier, R., Meyerhans, A., Roelants, G. & Wain-Hobson, S. (1990). Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature* **345,** 356-359.

Hunt, G. & Tiemessen, C.T. (2000). Occurrence of additional NF-kappaB-binding motifs in the long terminal repeat region of South African HIV type 1 subtype C isolates. *AIDS Res. Hum. Retroviruses* **16,** 305-306.

Hunt, G.M., Johnson, D. & Tiemesse, C.T. (2001). Characterisation of the long terminal repeat regions of South African human immunodeficiency virus type 1 isolates. *Virus Genes* **23,** 27-34.

Ivanov, D., Kwak, Y.T., Nee, E., Guo, J., Garcia-Martinez, L.F. & Gaynor, R.B. (1999). Cyclin T1 domains involved in complex formation with Tat and TAR RNA are critical for tat-activation. *J. Mol. Biol.* **288,** 41-56.

Ivanov, D., Kwak, Y.T., Guo, J. & Gaynor, R.B. (2000). Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol. Cell. Biol.* **20,** 2970-2983.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. & Varmus, H.E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331,** 280-283.

Jaffar, S., Wilkins, A., Ngom, P.T., Sabally, S., Corrah, T., Bangali, J.E., Rolfe, M. & Whittle, H.C. (1997). Rate of decline of percentage CD4+ cells is faster in HIV-1 than in HIV-2 infection. *J. Acquir. Immune. Defic. Syndr. Hum. Retrovirol.* **16,** 327-332.

Jaffe, H.W., Bregman, D.J. & Selik, R.M. (1983). Acquired immune deficiency syndrome in the United States: the first 1,000 cases. *J. Infect. Dis.* **148,** 339-345.

Jakobovits, A., Smith, D.H., Jakobovits, E.B. & Capon, D.J. (1988). A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. *Mol. Cell. Biol.* **8,** 2555-2561.

Jaleco, A.C., Covas, M.J. & Victorino, R.M. (1994). Analysis of lymphocyte cell death and apoptosis in HIV-2 infected patients. *Clin. Exp. Immunol.* **98,** 185-189.

Janssens, W., Fransen, K., Peeters, M., Heyndrickx, L., Motte, J., Bedjabaga, L., Delaporte, E., Piot, P. & Van der Groen, G. (1994). Phylogenetic analysis of a new chimpanzee lentivirus SIVcpz-gab2 from a wild-captured chimpanzee from Gabon. *AIDS. Res. Hum. Retroviruses* **10,** 1191-1192.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B. & Smale, S.T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol. Cell Biol.* **14,** 116-127.

Jeeninga, R.E., Hoogenkamp, M., Armand-Ugon, M., de Baar, M., Verhoef, K., & Berkhout, B. (2000). Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. *J. Virol.* **74,** 3740-3751.

Jhappan, C., Gallahan, D., Stahle, C., Chu, E., Smith, G.H., Merlino, G. & Callahan, R. (1992). Expression of an activated Notch-related int-3 transgene interferes with cell differentiation and induces neoplastic transformation in mammary and salivary glands. *Genes Dev.* **6,** 345-355.

John, S., Reeves, R.B., Lin, J.X., Child, R., Leiden, J.M., Thompson, C.B. & Leonard, W.J. (1995). Regulation of cell-type-specific interleukin-2 receptor alpha-chain gene expression: potential role of physical interactions between Elf-1, HMG-I(Y), and NF-kappa B family proteins. *Mol. Cell. Biol.* **15,** 1786-1796.

Johnson, P.R., Hamm, T.E., Goldstein, S., Kitov, S. & Hirsch, V.M. (1991). The genetic fate of molecularly cloned simian immunodeficiency virus in experimentally infected macaques. *Virology* **185,** 217-228.

Jones, K.A., Luciw, P.A. & Duchange, N. (1988). Structural arrangements of transcription control domains within the 5'-untranslated leader regions of the HIV-1 and HIV-2 promoters. *Genes Dev.* **2,** 1101-1114.

Jones, K.A. & Peterlin, B.M. (1994). Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* **63,** 717-743.

Jordan, A., Defechereux, P. & Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J.* **20,** 1726-1738.

Kadonaga, J.T. (1998). Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92,** 307-313.

Kalyanaraman, V.S., Sarngadharan, M.G., Robert-Guroff, M., Miyoshi, I., Golde, D. & Gallo, R.C. (1982). A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. *Science* **218,** 571-573.

Kanki, P.J., M'Boup, S., Ricard, D., Barin, F., Denis, F., Boye, C., Sangare, L., Travers, K., Albaum, M., Marlink, R., et al. (1987). Human T-lymphotropic virus type 4 and the human immunodeficiency virus in West Africa. *Science* **236,** 827-831.

Kanki, P.J., Travers, K.U., MBoup, S., Hsieh, C.C., Marlink, R.G., Gueye, N.A., Siby, T., Thior, I., Hernandez-Avila, M., Sankale, J.L., Ndoye, I. & Essex, M.E. (1994). Slower heterosexual spread of HIV-2 than HIV-1. *Lancet* **343,** 943-946.

Kao, S.Y., Calman, A.F., Luciw, P.A. & Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* **330,** 489-493.

Karlsson, G.B., Halloran, M., Li, J., Park, I.W., Gomila, R., Reimann, K.A., Axthelm, M.K., Iliff, S.A., Letvin, N.L. & Sodroski, J. (1997). Characterization of molecularly cloned simian-human immunodeficiency viruses causing rapid CD4+ lymphocyte depletion in rhesus monkeys. *J. Virol.* **71,** 4218-4225.

Kaslow, R.A., Duquesnoy, R., VanRaden, M., Kingsley, L., Marrari, M., Friedman, H., Su, S., Saah, A.J., Detels, R., Phair, J., et al. (1990). A1, Cw7, B8, DR3 HLA antigen combination associated with rapid

decline of T-helper lymphocytes in HIV-1 infection. A report from the Multicenter AIDS Cohort Study. *Lancet* **335,** 927-930.

Kaslow, R.A., Carrington, M., Apple, R., Park, L., Munoz, A., Saah, A.J., Goedert, J.J., Winkler, C., O'Brien, S.J., Rinaldo, C., Detels, R., Blattner, W., Phair, J., Erlich, H. & Mann, D.L. (1996). Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nat Med* **2,** 405-411.

Kassim, S., Wiktor, S.Z., Maurice, C., Lawson, A. & Coulibaly, I.M. (1998). Prevalence of HIV and syphilis infections among pregnant women attending urban antenatal clinics in Cote d'Ivoire, 1997. *12th World AIDS Conference, Geneva. [Abstract 13111].*

Kato, H., Sumimoto, H., Pognonec, P., Chen, C.H., Rosen, C.A. & Roeder, R.G. (1992). HIV-1 Tat acts as a processivity factor in vitro in conjunction with cellular elongation factors. *Genes Dev.* **6,** 655-666.

Katz, R.A., Merkel, G., Kulkosky, J., Leis, J. & Skalka, A.M. (1990). The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell* **63,** 87-95.

Kawakami, K., Scheidereit, C. & Roeder, R.G. (1988). Identification and purification of a human immunoglobulin-enhancer-binding protein (NF-kappa B) that activates transcription from a human immunodeficiency virus type 1 promoter in vitro. *Proc. Natl. Acad. Sci. USA.* **85,** 4700-4704.

Keet, I.P., Krijnen, P., Koot, M., Lange, J.M., Miedema, F., Goudsmit, J. & Coutinho, R.A. (1993). Predictors of rapid progression to AIDS in HIV-1 seroconverters. *AIDS* **7,** 51-57.

Keen, N.J., Churcher, M.J. & Karn, J. (1997). Transfer of Tat and release of TAR RNA during the activation of the human immunodeficiency virus type-1 transcription elongation complex. *EMBO J.* **16,** 5260-5272.

Kessler, M. & Mathews, M.B. (1992). Premature termination and processing of human immunodeficiency virus type 1-promoted transcripts. *J. Virol.* **66,** 4488-4496.

Kestler, H.W. 3rd, Ringler, D.J., Mori, K., Panicali, D.L., Sehgal, P.K., Daniel, M.D. & Desrosiers, R.C. (1991). Importance of the nef gene for maintenance of high virus loads and for development of AIDS. *Cell* **65,** 651-662.

Kiernan, R.E., Vanhulle, C., Schiltz, L., Adam, E., Xiao, H., Maudoux, F., Calomme, C., Burny, A., Nakatani, Y., Jeang, K.T., Benkirane, M. & Van Lint, C. (1999). HIV-1 tat transcriptional activity is regulated by acetylation. *EMBO J.* **18,** 6106-6118.

Killewo, J., Nyamuryekunge, K., Sandstrom, A., Bredberg-Raden, U., Wall, S., Mhalu, F. & Biberfeld, G. (1990). Prevalence of HIV-1 infection in the Kagera region of Tanzania: a population-based study. *AIDS* **4,** 1081-1085.

Kim, J.B. & Sharp, P.A. (2001). Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J. Biol. Chem.* **276,** 12317-12323.

Kim, J.L., Nikolov, D.B. & Burley, S.K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365,** 520-527.

Kim, S., Ikeuchi, K., Byrn, R., Groopman, J. & Baltimore, D. (1989). Lack of a negative influence on viral growth by the nef gene of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA.* **86,** 9544-9548.

Kim, T.K., Lagrange, T., Wang, Y.H., Griffith, J.D., Reinberg, D. & Ebright, R.H. (1997). Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proc. Natl. Acad. Sci. USA.* **94,** 12268-12273.

Kim, Y.K., Bourgeois, C.F., Isel, C., Churcher, M.J. & Karn, J. (2002). Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type 1 Tat-activated transcriptional elongation. *Mol. Cell. Biol.* **22,** 4622-4637.

Kimata, J.T., Kuller, L., Anderson, D.B., Dailey, P. & Overbaugh, J. (1999). Emerging cytopathic and antigenic simian immunodeficiency virus variants influence AIDS progression. *Nat. Med.* **5,** 535-541.

Kino, T., Slobodskaya, O., Pavlakis, G.N. & Chrousos, G.P. (2002). Nuclear receptor coactivator p160 proteins enhance the HIV-1 long terminal repeat promoter by bridging promoter-bound factors and the Tat-P-TEFb complex. *J. Biol. Chem.* **277,** 2396-2405.

Kirchhoff, F., Greenough, T.C., Brettler, D.B., Sullivan, J.L. & Desrosiers, R.C. (1995). Brief report: absence of intact nef sequences in a long-term survivor with nonprogressive HIV-1 infection. *N. Engl. J. Med.* **332,** 228-232.

Kirchoff, F., Greenough, T.C., Hamacher, M., Sullivan, J.L., & Desrosiers, R.C. (1997). Activity of human immunodeficiency virus type 1 promoter/TAR regions and *tat1* genes derived from individuals with different rates of disease progression. *Virology* **232,** 319-331.

Klasens, B.I., Das, A.T. & Berkhout, B. (1998). Inhibition of polyadenylation by stable RNA secondary structure. *Nucleic Acids Res.* **26,** 1870-1876.

Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.C. & Montagnier, L. (1984a). T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312,** 767-768.

Klatzmann, D., Barre-Sinoussi, F., Nugeyre, M.T., Danquet, C., Vilmer, E., Griscelli, C., Brun-Veziret, F., Rouzioux, C., Gluckman, J.C., Chermann, J.C., et al. (1984). Selective tropism of lymphadenopathy associated virus (LAV) for helper-inducer T lymphocytes. *Science* **225,** 59-63.

Klaver, B. & Berkhout, B. (1994). Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. *J. Virol.* **68,** 3830-3840.

Koken, S.E., van Wamel, J.L., Gouldsmit, J., Berkhout, B., & Geelen, J.L. (1992). Natural variants of the HIV-1 long terminal repeat. Analysis of promoters with duplicated DNA regulatory motifs. *Virology* **191,** 968-972.

Koken, S.E., Greijer, A.E., Verhoef, K., van Wamel, J., Bukrinskaya, A.G. & Berkhout, B. (1994). Intracellular analysis of in vitro modified HIV Tat protein. *J. Biol. Chem.* **269,** 8366-8375.

Kokkotou, E.G., Sankale, J.L., Mani, I., Gueye-Ndiaye, A., Schwartz, D., Essex, M.E., Mboup, S. & Kanki, P.J. (2000). In vitro correlates of HIV-2-mediated HIV-1 protection. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 6797-6802.

Koleske, A.J. & Young, R.A. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.* **20,** 113-116.

Kostrikis, L.G., Bagdades, E., Cao, Y., Zhang, L., Dimitriou, D. & Ho, D.D. (1995). Genetic analysis of human immunodeficiency virus type 1 strains from patients in Cyprus: identification of a new subtype designated subtype I. *J. Virol.* **69,** 6122-6130.

Kouinche Matagne, A., Tapko, J.B.T., Kamdem, T.K., Zokong, B.Z., Tookam, M.T. & Kaptue, L.N.K. (1998). *12th World AIDS Conference, Geneva. [Abstract 13126].*

Krebs, F.C., Mehrens, D., Pomeroy, S., Goodenow, M.M., & Wigdahl, B. (1998). Human immunodeficiency virus type 1 long terminal repeat quasispecies differ in basal transcription and nuclear factor recruitment in human glial cells and lymphocytes. *J. Biomed. Sci.* **5,** 31-44

Kuhnel, H., Von Briesen, H., Dietrich, U., Adamski, M., Mix, D., Biesert, L., Kreutz, R., Immelmann, A., Henco, K., Meichsner, C., et al. (1989). Molecular cloning of two west African human immunodeficiency virus type 2 isolates that replicate well in macrophages: a Gambian isolate, from a patient with neurologic acquired immunodeficiency syndrome, and a highly divergent Ghanian isolate. *Proc. Natl. Acad. Sci. USA.* **86,** 2383-2387.

Kwak, Y.T., Ivanov, D., Guo, J., Nee, E. & Gaynor, R.B. (1999). Role of the human and murine cyclin T proteins in regulating HIV-1 tat-activation. *J. Mol. Biol.* **288,** 57-69.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. & Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393,** 648-659.

Lagrange, T., Kim, T.K., Orphanides, G., Ebright, Y.W., Ebright, R.H. & Reinberg, D. (1996). High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 10620-10625.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D. & Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* **12,** 34-44.

Landau, N.R., Warton, M. & Littman, D.R. (1988). The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4. *Nature* **334,** 159-162.

Langley, R.J, Hirsch, V.M., O'Brien, S.J., Adger-Johnson, D., Goeken, R.M. & Olmsted, R.A. (1994). Nucleotide sequence analysis of puma lentivirus (PLV-14): genomic organization and relationship to other lentiviruses. *Virology* **202,** 853-864.

Langley, C.L., Benga-De, E., Critchlow, C.W., Ndoye, I., Mbengue-Ly, M.D., Kuypers, J., Woto-Gaye, G., Mboup, S., Bergeron, C., Holmes, K.K. & Kiviat, N.B. (1996). HIV-1, HIV-2, human papillomavirus infection and cervical neoplasia in high-risk African women. *AIDS* **10,** 413-417.

Lapham, C.K., Ouyang, J., Chandrasekhar, B., Nguyen, N.Y., Dimitrov, D.S. & Golding, H. (1996). Evidence for cell-surface association between fusin and the CD4-gp120 complex in human cell lines. *Science* **274,** 602-605.

Laspia, M.F., Wendel, P. & Mathews, M.B. (1993). HIV-1 Tat overcomes inefficient transcriptional elongation in vitro. *J. Mol. Biol.* **232,** 732-746.

Laybourn, P.J. & Dahmus, M.E. (1989). Transcription-dependent structural changes in the C-terminal domain of mammalian RNA polymerase subunit IIa/o. *J. Biol. Chem.* **264,** 6693-6698.

Learmont, J.C., Geczy, A.F., Mills, J., Ashton, L.J., Raynes-Greenow, C.H., Garsia, R.J., Dyer, W.B., McIntyre, L., Oelrichs, R.B., Rhodes, D.I., Deacon, N.J. & Sullivan, J.S. (1999). Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1. A report from the Sydney Blood Bank Cohort. *N. Engl. J. Med.* **340,** 1715-1722.

Lei, L., Ren, D. & Burton, Z.F. (1999). The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. *Mol. Cell. Biol.* **19,** 8372-8382.

Leiden, J.M., Wang, C.Y., Petryniak, B., Markovitz, D.M., Nabel, G.J. & Thompson, C.B. (1992). A novel Ets-related transcription factor, Elf-1, binds to human immunodeficiency virus type 2 regulatory elements that are required for inducible *trans* activation in T cells. *J. Virol.* **66,** 5890-5897.

Leitner, T., Alaeus, A., Marquina, S., Lilja, E., Lidman, K. & Albert, J. (1995). Yet another subtype of HIV type 1? *AIDS Res. Hum. Retroviruses* **11,** 995-997.

Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N. & Gregory, T.J. (1990). Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J. Biol. Chem.* **265,** 10373-10382.

Levy, J.A. (1998). HIV and the pathogenesis of AIDS, Second edn. Washington DC: American Society for Microbiology.

Levy, J.A., Hoffman, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M. & Oshiro, L.S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225,** 840-842.

Lewis, J.D. & Izaurralde, E. (1997). The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem.* **247,** 461-469.

Li, J.P. & Baltimore, D. (1991). Mechanism of leukemogenesis induced by mink cell focus-forming murine leukemia viruses. *J. Virol.* **65,** 2408-2414.

Liao, F., Alkhatib, G., Peden, K.W., Sharma, G., Berger, E.A. & Farber, J.M. (1997). STRL33, A novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1. *J. Exp. Med.* **185,** 2015-2023.

Lin, G., Lee, B., Haggarty, B.S., Doms, R.W. & Hoxie, J.A. (2001). CD4-independent use of Rhesus CCR5 by human immunodeficiency virus Type 2 implicates an electrostatic interaction between the CCR5 N terminus and the gp120 C4 domain. *J. Virol.* **75,** 10766-10778.

Lisse, I.M., Poulsen, A.G., Aaby, P., Knudsen, K. & Dias, F. (1996). Serial CD4 and CD8 T-lymphocyte counts and associated mortality in an HIV-2-infected population in Guinea-Bissau. *J. Acquir. Immune. Defic. Syndr. Hum. Retrovirol.* **13,** 355-362.

Littlefield, O., Korkhin, Y. & Sigler, P.B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc. Natl. Acad. Sci. USA.* **96,** 13668-13673.

Liu, Q., Gabriel, S.E., Roinick, K.L., Ward, R.D. & Arndt, K.M. (1999). Analysis of TFIIA function In vivo: evidence for a role in TATA-binding protein recruitment and gene-specific activation. *Mol. Cell. Biol.* **19,** 8673-8685.

Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A. & Landau, N.R. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86,** 367-377.

Liu, H.Y., Soda, Y., Shimizu, N., Haraguchi, Y., Jinno, A., Takeuchi, Y. & Hoshino, H. (2000). CD4-Dependent and CD4-independent utilization of coreceptors by human immunodeficiency viruses type 2 and simian immunodeficiency viruses. *Virology* **278,** 276-288.

Loussert-Ajaka, I., Chaix, M.L., Korber, B., Letourneur, F., Gomas, E., Allen, E., Ly, T.D., Brun-Vezinet, F., Simon, F. & Saragosti, S. (1995). Variability of human immunodeficiency virus type 1 group O strains isolated from Cameroonian patients living in France. *J. Virol.* **69,** 5640-5649.

Louwagie, J., McCutchan, F.E., Peeters, M., Brennan, T.P., Sanders-Buell, E., Eddy, G.A., Van der Groen, G., Fransen, K., Gershy-Damet, G.M., Deleys, R., et al. (1993). Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS* **7,** 769-780.

Lukashov, V.V., Kuiken, C.L. & Goudsmit, J. (1995). Intrahost human immunodeficiency virus type 1 evolution is related to length of the immunocompetent period. *J. Virol.* **69,** 6911-6916.

Mabey, D.C., Tedder, R.S., Hughes, A.S., Corrah, P.T., Goodison, S.J., O'Connor, T., Shenton, F.C., Lucas, S.B., Whittle, H.C. & Greenwood, B.M. (1988). Human retroviral infections in The Gambia: prevalence and clinical features. *Br. Med. J. (Clin. Res. Ed.)* **296,** 83-86.

Madani, N. & Kabat, D. (1998). An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. *J. Virol.* **72,** 10251-10255.

Maddon, P.J., McDougal, J.S., Clapham, P.R., Dalgleish, A.G., Jamal, S., Weiss, R.A. & Axel, R. (1988). HIV infection does not require endocytosis of its receptor, CD4. *Cell* **54,** 865-874.

Madore, S.J. & Cullen, B.R. (1993). Genetic analysis of the cofactor requirement for human immunodeficiency virus type 1 Tat function. *J. Virol.* **67,** 3703-3711.

Magierowska, M., Theodorou, I., Debre, P., Sanson, F., Autran, B., Riviere, Y., Charron, D. & Costagliola, D. (1999). Combined genotypes of CCR5, CCR2, SDF1, and HLA genes can predict the long-term nonprogressor status in human immunodeficiency virus-1-infected individuals. *Blood* **93,** 936-941.

Malik, A., et al., (unpublished).

Mancebo, H.S., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D. & Flores, O. (1997). P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. *Genes Dev.* **11,** 2633-2644.

Mangasarian, A., Piguet, V., Wang, J.K., Chen, Y.L. & Trono, D. (1999). Nef-induced CD4 and major histocompatibility complex class I (MHC-I) down-regulation are governed by distinct determinants: N-terminal alpha helix and proline repeat of Nef selectively regulate MHC-I trafficking. *J. Virol.* **73,** 1964-1973.

Mansfield, K.G., Lerch, N.W., Gardner, M.B. & Lackner, A.A. (1995). Origins of simian immunodeficiency virus infection in macaques at the New England Regional Primate Research Center. *J Med Primatol* **24,** 116-122.

Marcello, A., Zoppe, M. & Giacca, M. (2001). Multiple modes of transcriptional regulation by the HIV-1 Tat transactivator. *IUBMB Life* **51,** 175-181.

Marciniak, R.A. & Sharp, P.A. (1991). HIV-1 Tat protein promotes formation of more-processive elongation complexes. *EMBO J.* **10,** 4189-4196.

Mariani, R., Kirchhoff, F., Greenough, T.C., Sullivan, J.L., Desrosiers, R.C. & Skowronski, J. (1996). High frequency of defective nef alleles in a long-term survivor with nonprogressive human immunodeficiency virus type 1 infection. *J. Virol.* **70,** 7752-7764.

Markovitz, D.M., Hannibal, M., Perez, V.L., Gaunti, C., Folks, T.M., Nabel, G.J. (1990). Differential regulation of human immunodeficiency viruses (HIVs): a specific regulatory element in HIV-2 responds to stimulation of the T-cell antigen receptor. *Proc. Natl. Acad. Sci. USA.* **87,** 9098-9102.

Markovitz, D.M., Smith, M.J., Hilfinger, J., Hannibal, M.C., Petryniak, B. & Nabel, G.J. (1992). Activation of the human immunodeficiency virus type 2 enhancer is dependent on purine box and κB regulatory elements. *J. Virol.* **66,** 5479-5484.

Marlink, R., Kanki, P., Thior, I., Travers, K., Eisen, G., Siby, T., Traore, I., Hsieh, C.C., Dia, M.C., Gueye, E.H., et al. (1994). Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science* **265,** 1587-1590.

Marshall, N.F. & Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J. Biol. Chem.* **270,** 12335-12338.

Marshall, N.F., Peng, J., Xie, Z. & Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J. Biol. Chem.* **271,** 27176-27183.

Martin, M. & Freed, E. (1996). Structure/Function studies of the HIV-1 MA protein. *Cell biology of Virus Entry, Replication and Pathogenesis. Keystone Symposia on Molecular and Cellular biology.* **Sante Fe, New Mexico.**

Marx, P.A., Munn, R.J. & Joy, K.I. (1988). Computer emulation of thin section electron microscopy predicts an envelope-associated icosadeltahedral capsid for human immunodeficiency virus. *Lab. Invest.* **58,** 112-118.

Marzio, G., Tyagi, M., Gutierrez, M.I. & Giacca, M. (1998). HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 13519-13524.

Matheron, S., Mendoza-Sassi, G., Simon, F., Olivares, R., Coulaud, J.P. Brun-Vezinet, F. (1997). HIV-1 and HIV-2 AIDS in African patients living in Paris. *AIDS* **11,** 934-936.

Maury, W. (1994). Monocyte maturation controls expression of equine infectious anemia virus. *J. Virol.* **68,** 6270-6279.

Maury, W., Perryman, S., Oaks, J.L., Seid, B.K., Crawford, T., McGuire, T. & Carpenter, S. (1997). Localized sequence heterogeneity in the long terminal repeats of in vivo isolates of equine infectious anemia virus. *J. Virol.* **71,** 4929-4937.

Maury, W. (1998). Regulation of equine infectious anemia virus expression. *J Biomed Sci.* **5,** 11-23.

Maury, W., Bradley, S., Wright, B. & Hines, R. (2000). Cell specificity of the transcription-factor repertoire used by a lentivirus: motifs important for expression of equine infectious anemia virus in nonmonocytic cells. *Virology* **267,** 267-278.

McAdam, S., Kaleebu, P., Krausa, P., Goulder, P., French, N., Collin, B., Blanchard, T., Whitworth, J., McMichael, A. & Gotch F. (1998). Cross-clade recognition of p55 by cytotoxic T lymphocytes in HIV-1 infection. *AIDS* **12,** 571-579.

McAllister, J.J., Phillips, D., Millhouse, S., Conner, J., Hogan, T., Ross, H.L. & Wigdahl, B. (2000). Analysis of the HIV-1 LTR NF-kappaB-proximal Sp site III: evidence for cell type-specific gene regulation and viral replication. *Virology* **274,** 262-277.

McBride, M.S. & Panganiban, A.T. (1997). Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J. Virol.* **71,** 2050-2058.

McBride, M.S., Schwartz, M.D. & Panganiban, A.T. (1997). Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. *J. Virol.* **71,** 4544-4554.

McKnight, A., Dittmar, M.T., Moniz-Periera, J., Ariyoshi, K., Reeves, J.D., Hibbitts, S., Whitby, D., Aarons, E., Proudfoot, A.E., Whittle, H. & Clapham, P.R. (1998). A broad range of chemokine receptors are used by primary isolates of human immunodeficiency virus type 2 as coreceptors with CD4. *J. Virol.* **72,** 4065-4071.

McNearney, T., Hornickova, Z., Templeton, A., Birdwell, A., Arens, M., Markham, R., Saah, A. & Ratner, L. (1995). Nef and LTR sequence variation from sequentially derived human immunodeficiency virus type 1 isolates. *Virology* **208,** 388-398.

Mellors, J.W., Rinaldo, C.R.Jr, Gupta, P., White, R.M., Todd, J.A. & Kingsley, L.A. (1996). Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science* **272,** 1167-1170.

Mellors, J.W., Munoz, A., Giorgi, J.V., Margolick, J.B., Tassoni, C.J., Gupta, P., Kingsley, L.A., Todd, J.A., Saah, A.J., Detels, R., Phair, J.P. & Rinaldo, C.R.Jr. (1997). Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann. Intern. Med.* **126,** 946-954.

Meyerhans, A., Cheynier, R., Albert, J., Seth, M., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B. & Wain-Hobson, S. (1989). Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58,** 901-910.

Michael, N.L., Vahey, M., Burke, D.S. & Redfield, R.R. (1992). Viral DNA and mRNA expression correlate with the stage of human immunodeficiency virus (HIV) type 1 infection in humans: evidence for viral replication in all stages of HIV disease. *J. Virol.* **66,** 310-316.

Micheal, N.L., D'Arcy, L., Ehrenberg, P.K., & Redfield, R.R. (1994). Naturally occurring genotypes of the human immunodeficiency virus type 1 long terminal repeat display a wide range of basal and Tat-induced transcriptional activities. *J. Virol.* **68,** 3163-3174.

Michel, P., Balde, A.T., Roussilhon, C., Aribot, G., Sarthou, J.L., Gougeon, M.L. (2000). Reduced immune activation and T cell apoptosis in human immunodeficiency virus type 2 compared with type 1: correlation of T cell apoptosis with β_2 microglobulin concentration and disease evolution. *J. Infect. Dis.* **181,** 64-75.

Mildvan, D., Mathur, U., Enlow, R.W., Romain, P.L., Winchester, R.J., Colp, C., Singman, H., Adelsberg, B.R. & Spigland, I. (1982). Opportunistic infections and immune deficiency in homosexual men. *Ann. Intern. Med.* **96,** 700-704.

Miller, M., Kennewell, A., Takayama, Y., Bruskin, A., Bishop, J.M., Johnson, G. & Symonds, G. (1990). Transformation of early erythroid precursor cells (BFU-E) by a recombinant murine retrovirus containing v-erb-B. *Oncogene* **5,** 1125-31.

Miura, T., Sakuragi, J., Kawamura, M., Fukasawa, M., Moriyama, E.N., Gojobori, T., Ishikawa, K., Mingle, J.A., Nettey, V.B., Akari, H., et al. (1990). Establishment of a phylogenetic survey system for AIDS-related lentiviruses and demonstration of a new HIV-2 subgroup. *AIDS* **4,** 1257-1261.

Modrow. S., Hahn, B.H., Shaw, G.M., Gallo, R.C., Wong-Staal, F. & Wolf, H. (1987). Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *J. Virol.* **61,** 570-578.

Montano, M.A., Novitsky, V.A., Blackard, J.T., Cho, N.L., Katzenstein, D.A., & Essex, M. (1997). Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *J. Virol.* **71,** 8657-8665.

Montano, M.A., Nixon, C.P., & Essex, M. (1998). Dysregulation through the NF-κB enhancer and TATA box of the human immunodeficiency virus type 1 subtype E promoter. *J. Virol.* **72,** 8446-8452.

Moore, J.P. (1993). Viral fusion mechanisms. Edited by J. Bentz: CRC Press.

Morgan, D., Maude, G.H., Malamba, S.S., Okongo, M.J., Wagner, H.U., Mulder, D.W. & Whitworth, J.A. (1997). HIV-1 disease progression and AIDS-defining disorders in rural Uganda. *Lancet* **350,** 245-250.

Morris, D.W. (1991). Molecular biology and pathogenesis of mouse mammary tumour virus. *Rev. Med. Virol.* **1,** 223-233.

Moses, A.V., Ibanez, C., Gaynor, R., Ghazal, P. & Nelson, J.A. (1994). Differential role of long terminal repeat control elements for the regulation of basal and Tat-mediated transcription of the human immunodeficiency virus in stimulated and unstimulated primary human macrophages. *J. Virol.* **68,** 298-307.

Mulanga-Kabeya, C., Nzilambi, N., Edidi, B., Minlangu, M., Tshimpaka, T., Kambembo, L., Atibu, L., Mama, N., Ilunga, W., Sema, H., Tshimanga, K., Bongo, B., Peeters, M. & Delaporte, E. (1998). Evidence of stable HIV seroprevalences in selected populations in the Democratic Republic of the Congo. *AIDS* **12,** 905-910.

Mullis, K.B. & Faloona, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods. Enzymol.* **155,** 335-350.

Murphey-Corb, M., Martin, L.N., Rangan, S.R., Baskin, G.B., Gormus, B.J., Wolf, R.H., Andes, W.A., West, M. & Montelaro, R.C. (1986). Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. *Nature* **321,** 435-437.

Myers, G., Rabson, A.B., Berzofsky, J.A., Smith, T.F. & Wong-Staal, F. (1990). Human retroviruses and AIDS. A compilation and analysis of nucleic acid and amino acid sequences. *Theoretical Biology and Biophysics.*

Myers, G. & Pavlakis, G.N. (1992). Evolutionary potential of complex retroviruses. In *The Retroviridae*, pp 51-105. Edited by J.A.Levy. New York.: Plenum Press.

Myers, G., Korber, B., Wain-Hobson, S., Smith, R.F. & Pavlakis, G.N. (1993). Human Retroviruses and AIDS. Los Alonos: Los Alomos National Laboratory.

Nabel, G. & Baltimore, D. (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* **326,** 711-713.

Naghavi, M.H., Salminen, M.O., Sonnerborg, A. & Vahlne, A. (1999). DNA sequence of the long terminal repeat of human immunodeficiency virus type 1 subtype A through G. *AIDS Res. Hum. Retroviruses* **15,**485-488.

Narayan, O. & Clements, J.E. (1989). Biology and pathogenesis of lentiviruses. *J. Gen. Virol.* **70,** 1617-1639.

Naucler, A., Andreasson, P.A., Costa, C.M., Thorstensson, R. & Biberfeld, G. (1989). HIV-2-associated AIDS and HIV-1 seroprevalence in Bissau, Guinea-Bissau. *J.Acquir.Immune.Defic.Syndr.* **2,** 88-93.

Neel, B.G., Hayward, W.S., Robinson, H.L., Fang, J. & Astrin, S.M. (1981). Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. *Cell* **23,** 323-334.

Neil, J.C., Forrest, D., Doggett, D.L. & Mullins, J.I. (1987). The role of feline leukaemia virus in naturally occurring leukaemias. *Cancer. Surv.* **6,** 117-137.

Neil, J.C., Fulton, R., Rigby, M. & Stewart, M. (1991). Feline leukaemia virus: generation of pathogenic and oncogenic variants. *Curr. Top. Microbiol. Immunol.* **171,** 67-93.

Nermut, M.V., Grief, C., Hashmi, S., Hockley, D.J. (1993). Further evidence of icosahedral symmetry in human and simian immunodeficiency virus. *AIDS Res. Hum. Retroviruses* **9,** 929-938.

Neugebauer, K.M. & Roth, M.B. (1997). Transcription units as RNA processing units. *Genes Dev.* **11,** 3279-3285.

Neville, M., Stutz, F., Lee, L., Davis, L.I. & Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* **7,** 767-775.

Niederman, T.M., Thielan, B.J. & Ratner, L. (1989). Human immunodeficiency virus type 1 negative factor is a transcriptional silencer. *Proc. Natl. Acad. Sci. USA.* **86,** 1128-1132.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G. & Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377,** 119-128.

Nolan, G.P., Ghosh, S., Liou, H.C., Tempst, P. & Baltimore, D. (1991). DNA binding and I kappa B inhibition of the cloned p65 subunit of NF-kappa B, a rel-related polypeptide. *Cell* **64,** 961-969.

Norrgren, H., Andersson, S., Naucler, A., Dias, F., Johansson, I. & Biberfeld, G. (1995). HIV-1, HIV-2, HTLV-I/II and Treponema pallidum infections: incidence, prevalence, and HIV-2-associated mortality in an occupational cohort in Guinea-Bissau. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **9,** 422-428.

Norrgren, H., Andersson, S., Dias, F., Naucler, A. & Biberfeld, G. (1997). Trends of incidence and prevalence of HIV-1 and HIV-2 in Guinea-Bissau, West Africa. *Xth International Conference on AIDS and STD in Africa, Abidjan. [Abstract B.045].*

Norrgren, H., Marquina, S., Leitner, T., Aaby, P., Melbye, M., Poulsen, A.G., Larsen, O., Dias, F., Escanilla, D., Andersson, S., Albert, J. & Naucler, A. (1997b). HIV-2 genetic variation and DNA load in asymptomatic carriers and AIDS cases in Guinea-Bissau. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **16,** 31-38.

Novembre, F.J., Hirsch, V.M., McClure, H.M., Fultz, P.N. & Johnson, P.R. (1992). SIV from stump-tailed macaques: molecular characterization of a highly transmissible primate lentivirus. *Virology* **186,** 783-787.

Novembre, F.J., Johnson, P.R., Lewis, M.G., Anderson, D.C., Klumpp, S., McClure, H.M. & Hirsch, V.M. (1993). Multiple viral determinants contribute to pathogenicity of the acutely lethal simian immunodeficiency virus SIVsmmPBj variant. *J. Virol.* **67,** 2466-2474.

Novitsky, V., Smith, U.R., Gilbert, P., McLane, M.F., Chigwedere, P., Williamson, C., Ndung'u, T., Klein, I., Chang, S.Y., Peter, T., Thior, I., Foley, B.T., Gaolekwe, S., Rybak, N., Gaseitsiwe, S., Vannberg, F., Marlink, R., Lee, T.H. & Essex, M. (2002). Human immunodeficiency virus type 1 subtype C molecular phylogeny: consensus sequence for an AIDS vaccine design? *J. Virol.* **76,** 5435-5451.

Nowak, M.A., Anderson, R.M., McLean, A.R., Wolfs, T.F.W., Goudsmit, J. & May, R.M. (1991). Antigenic Diversity Thresholds and the Development of AIDS. *Science* **254,** 963-969.

Nunn, A.J., Mulder, D.W., Kamali, A., Ruberantwari, A., Kengeya-Kayondo, J.F. & Whitworth, J. (1997). Mortality associated with HIV-1 infection over five years in a rural Ugandan population: cohort study. *B.M.J.* **315,** 767-771.

Nusse, R., Theunissen, H., Wagenaar, E., Rijsewijk, F., Gennissen, A., Otte, A., Schuuring, E. & van Ooyen, A. (1990). The Wnt-1 (int-1) oncogene promoter and its mechanism of activation by insertion of proviral DNA of the mouse mammary tumor virus. *Mol. Cell. Biol.* **10,** 4170-4179.

Nusse, R. (1991). Insertional mutagenesis in mouse mammary tumorigenesis. *Curr. Top. Microbiol. Immunol.* **171,** 43-65.

Oberlin, E., Amara, A., Bachelerie, F., Bessia, C., Virelizier, J.L., Arenzana-Seisdedos, F., Schwartz, O., Heard, J.M., Clark-Lewis, I., Legler, D.F., Loetscher, M., Baggiolini, M. & Moser, B. (1996). The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* **382,** 833-835.

Obisesan, K.A., Olaleye, O.D. & Adeyemo, A.A. (1997). The increasing prevalence of HIV-1 and HIV-2 infections in a low risk antenatal population in south west Nigeria. *Int. J. Gynaecol. Obstet.* **56,** 271-273.

O'Brien, T.R., George, J.R. & Holmberg, S.D. (1992). Human immunodeficiency virus type 2 infection in the United States. Epidemiology, diagnosis, and public health implications. *JAMA* **267,** 2775-2779.

O'Brien, W.A., Hartigan, P.M., Martin, D., Esinhart, J., Hill, A., Benoit, S., Rubin, M., Simberkoff, M.S. & Hamilton, J.D. (1996). Changes in plasma HIV-1 RNA and CD4+ lymphocyte counts and the risk of progression to AIDS. Veterans Affairs Cooperative Study Group on AIDS. *N. Engl. J. Med.* **334,** 426-431.

O'Donovan, D., Ariyoshi, K., Milligan, P., Ota, M., Yamuah, L., Sarge-Njie, R. & Whittle, H. (2000). Maternal plasma viral RNA levels determine marked differences in mother-to-child transmission rates of HIV-1 and HIV-2 in The Gambia. *AIDS* **14,** 441-448.

Okamoto, H., Sheline, C.T., Corden, J.L., Jones, K.A. & Peterlin, B.M. (1996). Trans-activation by human immunodeficiency virus Tat protein requires the C-terminal domain of RNA polymerase II. *Proc. Natl. Acad. Sci. USA.* **93,** 11575-11579.

Okarofor, L., et al., (unpublished).

Olmsted, R.A., Barnes, A.K., Yamamoto, J.K., Hirsch, V.M., Purcell, R.H. & Johnson, P.R. (1989a). Molecular cloning of feline immunodeficiency virus. *Proc. Natl. Acad. Sci. USA.* **86,** 2448-2452.

Olmsted, R.A., Hirsch, V.M., Purcell, R.H. & Johnson, P.R. (1989b). Nucleotide sequence analysis of feline immunodeficiency virus: genome organization and relationship to other lentiviruses. *Proc. Natl. Acad. Sci. USA.* **86,** 8088-8092.

Onuma, M., Koomoto, E., Furuyama, H., Yasutomi, Y., Taniyama, H., Iwai, H. & Kawakami, Y. (1992). Infection and dysfunction of monocytes induced by experimental inoculation of calves with bovine immunodeficiency-like virus. *J. Acquir. Immune Defic. Syndr.* **5,** 1009-1015.

Oroszlan, S. & Luftig, R.B. (1990). Retroviral Proteinases. *Curr. Top. Micro. Immunol.* **157,** 153-185.

Orphanides, G., Lagrange, T. & Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev.* **10,** 2657-2683.

Ou, S.H., Garcia-Martinez, L.F., Paulssen, E.J. & Gaynor, R.B. (1994). Role of flanking E box motifs in human immunodeficiency virus type 1 TATA element function. *J. Virol.* **68,** 7188-7199.

Owen, S.M., Ellenberger, D., Rayfield, M., Wiktor, S., Michel, P., Grieco, M.H., Gao, F., Hahn, B.H. & Lal, R.B. (1998). Genetically divergent strains of human immunodeficiency virus type 2 use multiple coreceptors for viral entry. *J. Virol.* **72,** 5425-5432.

Padua, E., Jenkins, A., Brown, S., Bootman, J., Paixao, M.T., Almond, N. & Berry, N. (2003). Natural variation of the nef gene in human immunodeficiency virus type 2 infections in Portugal. *J. Gen. Virol.* **84,** 1287-1299.

Pagtakhan, A.S. & Tong-Starksen, S.E. (1997). Interactions between Tat of HIV-2 and transcription factor Sp1. *Virology* **238,** 221-230.

Pantaleo, G., Demarest, J.F., Soudeyns, H., Graziosi, C., Denis, F., Adelsberger, J.W., Borrow, P., Saag, M.S., Shaw, G.M., Sekaly, R.P., et al. (1994). Major expansion of CD8+ T cells with a predominant V beta usage during the primary immune response to HIV. *Nature* **370,** 463-467.

Pantaleo, G., Graziosi, C., Demarest, J.F., Cohen, O.J., Vaccarezza, M., Gantt, K., Muro-Cacho, C. & Fauci, A.S. (1994). Role of lymphoid organs in the pathogenesis of human immunodeficiency virus (HIV) infection. *Immunol. Rev.* **140,** 105-130.

Parada, C.A. & Roeder, R.G. (1996). Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature* **384,** 375-378.

Parrott, C., Seidner, T., Duh, E., Leonard, J., Theodore, T.S., Buckler-White, A., Martin, M.A. & Rabson, A.B. (1991). Variable role of the long terminal repeat Sp1-binding sites in human immunodeficiency virus replication in T lymphocytes. *J. Virol.* **65,** 1414-1419.

Patikoglou, G.A., Kim, J.L., Sun, L., Yang, S.H., Kodadek, T. & Burley, S.K. (1999). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* **13,** 3217-3230.

Payne, G.S., Courtneidge, S.A., Crittenden, L.B., Fadly, A.M., Bishop, J.M. & Varmus, H.E. (1981). Analysis of avian leukosis virus DNA and RNA in bursal tumours: viral gene expression is not required for maintenance of the tumor state. *Cell* **23,** 311-322.

Payne, L.N. (1992). Biology of avian retroviruses. *The Retroviridae.* New York: Plenum Press, 482-483.

Payne, S.L., La Celle, K., Pei, X.F., Qi, X.M., Shao, H., Steagall, W.K., Perry, S. & Fuller, F. (1999). Long terminal repeat sequences of equine infectious anaemia virus are a major determinant of cell tropism. *J. Gen Virol.* **80,** 755-759.

Pedersen, N.C., Ho, E.W., Brown, M.L. & Yamamoto, J.K. (1987). Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. *Science* **235,** 790-793.

Pedersen, N.C., Yamamoto, J.K., Ishida, T. & Hansen, H. (1989). Feline immunodeficiency virus infection. *Vet Immunol Immunopathol* **21,** 111-129.

Peeters, M., Honore, C., Huet, T., Bedjabaga, L., Ossari, S., Bussi, P., Cooper, R.W. & Delaporte, E. (1989). Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. *AIDS* **3,** 625-630.

Peeters, M., Koumare, B., Mulanga, E., et al. (1998). Genetic subtypes of HIV type 1 and HIV type 2 strains in commercial sex workers from Bamako, Mali. *AIDS Res. Hum. Retroviruses* **14,** 51-58.

Peng, J., Zhu, Y., Milton, J.T. & Price, D.H. (1998). Identification of multiple cyclin subunits of human P-TEFb. *Genes Dev.* **12,** 755-762.

Pepin, J., Dunn, D., Gaye, I., Alonso, P., Egboga, A., Tedder, R., Piot, P., Berry, N., Schellenberg, D., Whittle, H. (1991a). HIV-2 infection among prostitutes working in The Gambia: association with serological evidence of genital ulcer diseases and with generalised lymphadenopathy. *AIDS* **5,** 69-75.

Pepin, J., Morgan, G., Dunn, D., Gevao, S., Mendy, M., Gaye, I., Scollen, N., Tedder, R. & Whittle, H. (1991b). HIV-2-induced immunosuppression among asymptomatic West African prostitutes: evidence that HIV-2 is pathogenic, but less so than HIV-1. *AIDS* **5,** 1165-1172.

Peterson, A. & Seed, B. (1988). Genetic analysis of monoclonal antibody and HIV binding sites on the human lymphocyte antigen CD4. *Cell* **54,** 65-72.

Pickering, H., Todd, J., Dunn, D., Pepin, J. & Wilkins, A. (1992). Prostitutes and their clients: a Gambian survey. *Soc. Sci. Med.* **34,** 75-88.

Pieniazek, D., Peralta, J.M., Ferreira, J.A., Krebs, J.W., Owen, S.M., Sion, F.S., Filho, C.F., Sereno, A.B., de Sa, C.A., Weniger, B.G., et al. (1991). Identification of mixed HIV-1/HIV-2 infections in Brazil by polymerase chain reaction. *AIDS* **5,** 1293-1299.

Piguet, V. & Trono, D. (1999). The Nef protein of primate lentiviruses. *Reviews in Medical Virology* **9,** 111-120.

Ping, Y.H. & Rana, T.M. (1999). Tat-associated kinase (P-TEFb): a component of transcription preinitiation and elongation complexes. *J. Biol. Chem.* **274,** 7399-7404.

Ping, Y.H. & Rana, T.M. (2001). DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J. Biol. Chem.* **276,** 12951-12958.

Poeschla, E.M., Wong-Staal, F. & Looney, D.J. (1998). Efficient transduction of nondividing human cells by feline immunodeficiency virus lentiviral vectors. *Nat. Med.* **4,** 354-357.

Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. & Gallo, R.C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. USA.* **77,** 7415-7419.

Popper, S.J., Sarr, A.D., Travers, K.U., Gueye-Ndiaye, A., Mboup, S., Essex, M.E. & Kanki, P.J. (1999). Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *J. Infect. Dis.* **180,** 1116-1121.

Popper, S.J., Sarr, A.D., Gueye-Ndiaye, A., Mboup, S., Essex, M.E. & Kanki, P.J. (2000). Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in vivo. *J. Virol.* **74,** 1554-1557.

Poulsen, A.G., Kvinesdal, B., Aaby, P., Molbak, K., Frederiksen, K., Dias, F. & Lauritzen, E. (1989). Prevalence of and mortality from human immunodeficiency virus type 2 in Bissau, West Africa. *Lancet* **1,** 827-831.

Poulsen, A.G., Aaby, P., Gottschau, A., Kvinesdal, B.B., Dias, F., Molbak, K. & Lauritzen, E. (1993). HIV-2 infection in Bissau, West Africa, 1987-1989: incidence, prevalences, and routes of transmission. *J.Acquir.Immune.Defic.Syndr* **6,** 941-948.

Poulsen, A.G., Aaby, P., Larsen, O., Jensen, H., Naucler, A., Lisse, I.M., Christiansen, C.B., Dias, F. & Melbye, M. (1997). 9-year HIV-2-associated mortality in an urban community in Bissau, West Africa. *Lancet* **349,** 911-914.

Preston, B.D., Poiesz, B.J. & Loeb, L.A. (1988). Fidelity of HIV-1 reverse transcriptase. *Science* **242,** 1168-1171.

Profy, A.T., Salinas, P.A., Eckler, L.I., Dunlop, N.M., Nara, P.L. & Putney, S.D. (1990). Epitopes recognized by the neutralizing antibodies of an HIV-1-infected individual. *J. Immunol.* **144,** 4641-4647.

Quinn, T.C. (1994). Population migration and the spread of types 1 and 2 human immunodeficiency viruses. *Proc. Natl. Acad. Sci. USA.* **91,** 2407-2414.

Quiñones-Mateu, M.E., Mas, A., Lain de Lera, T., Soriano, V., Alcami, J., Lederman, M.M., Domingo, E. (1998). LTR and *tat* variability of HIV-1 isolates from patients with divergent rates of disease progression. *Virus Res.* **57,** 11-20.

Ranish, J.A., Yudkovsky, N. & Hahn, S. (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* **13,** 49-63.

Reed-Inderbitzin, E. & Maury, W. (2003). Cellular specificity of HIV-1 replication can be controlled by LTR sequences. *Virology* **314,** 680-695.

Reeves, J.D., Hibbitts, S., Simmons, G., McKnight, A., Azevedo-Pereira J.M., Moniz-Pereira, J. & Clapham, P.R. (1999). Primary human immunodeficiency virus type 2 (HIV-2) isolates infect CD4-

negative cells via CCR5 and CXCR4: comparison with HIV-1 and simian immunodeficiency virus and relevance to cell tropism in vivo. *J. Virol.* **73,** 7795-7804.

Reeves, J.D. & Doms, R.W. (2002). Human immunodeficiency virus type 2. *J. Gen. Virol.* **83,** 1253-1265.

Reiss, P., de Wolf, F., Kuiken, C.L., de Ronde, A., Dekker, J., Boucher, C.A., Debouck, C., Lange, J.M. & Goudsmit, J. (1991). Contribution of antibody response to recombinant HIV-1 gene-encoded products nef, rev, tat, and protease in predicting development of AIDS in HIV-1-infected individuals. *J. Acquir. Immune Defic. Syndr.* **4,** 165-172.

Remy, G. (1998). HIV-2 infection throughout the world. A geographical perspective. *Santé* **8,** 440-446.

Renjifo, B., Speck, N.A., Winandy, S., Hopkins, N. & Li, Y. (1990). cis-acting elements in the U3 region of a simian immunodeficiency virus. *J. Virol.* **64,** 3130-3134.

Rethwilm, A., Erlwein, O., Baunach, G., Maurer, B. & ter Meulen, V. (1991). The transcriptional transactivator of human foamy virus maps to the bel 1 genomic region. *Proc. Natl. Acad. Sci. USA.* **88,** 941-945.

Rey, M.A., Girard, P.M., Harzic, M., Madjar, J.J., Brun-Vezinet, F. & Saimot, A.G. (1987). HIV-1 and HIV-2 double infection in French homosexual male with AIDS-related complex (Paris, 1985). *Lancet* **1,** 388-389.

Rhim, H. & Rice, A.P. (1993). TAR RNA binding properties and relative transactivation activities of human immunodeficiency virus type 1 and 2 Tat proteins. *J. Virol.* **67,** 1110-1121.

Rhim, H. & Rice, A.P. (1994a). Functional significance of the dinucleotide bulge in stem-loop1 and stem-loop2 of HIV-2 TAR RNA. *Virology* **202,** 202-211.

Rhim, H. & Rice, A.P. (1994b). Exon2 of HIV-2 Tat contributes to transactivation of the HIV-2 LTR by increasing binding affinity to HIV-2 TAR RNA. *Nucleic Acids Res.* **22,** 4405-4413.

Ricard, D., Wilkins, A., N'Gum, P.T., Hayes, R., Morgan, G., Da Silva, A.P. & Whittle, H. (1994). The effects of HIV-2 infection in a rural area of Guinea-Bissau. *AIDS* **8,** 977-982.

Richter, S., Cao, H. & Rana, T.M. (2002a). Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation. *Biochemistry* **41,** 6391-6397.

Richter, S., Ping, Y.H. & Rana, T.M. (2002b). TAR RNA loop: a scaffold for the assembly of a regulatory switch in HIV replication. *Proc. Natl. Acad. Sci. USA.* **99,** 7928-7933.

Rittner, K., Churcher, M.J., Gait, M.J. & Karn, J. (1995). The human immunodeficiency virus long terminal repeat includes a specialised initiator element which is required for Tat-responsive transcription. *J. Mol. Biol.* **248,** 562-580.

Roberts, J.D., Bebenek, K. & Kunkel, T.A. (1988). The accuracy of reverse transcriptase from HIV-1. *Science* **242,** 1171-1173.

Robert-Guroff, M., Aldrich, K., Muldoon, R., Stern, T.L., Bansal, G.P., Matthews, T.J., Markham, P.D., Gallo, R.C. & Franchini, G. (1992). Cross-neutralization of human immunodeficiency virus type 1 and 2 and simian immunodeficiency virus isolates. *J. Virol.* **66,** 3602-3608.

Robinson, H.L. & Gagnon, G.C. (1986). Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. *J. Virol.* **57,** 28-36.

Rodenburg, C.M., Li, Y., Trask, S.A., Chen, Y., Decker, J., Robertson, D.L., Kalish, M.L., Shaw, G.M., Allen, S., Hahn, B.H., Gao, F.; UNAIDS and NIAID Networks for HIV Isolation and Characterization. (2001). Near full-length clones and reference sequences for subtype C isolates of HIV type 1 from three different continents. *AIDS Res. Hum. Retroviruses* **17,** 161-168.

Rodes, B., Holguin, A., Soriano, V., Dourana, M., Mansinho, K., Antunes, F. & Gonzalez-Lahoz, J. (2000). Emergence of drug resistance mutations in human immunodeficiency virus type 2-infected subjects undergoing antiretroviral therapy. *J. Clin. Microbiol.* **38,** 1370-1374.

Roof, P., Ricci, M., Genin, P., Montano, M.A., Essex, M., Wainberg, M.A., Gatignol, A. & Hiscott, J. (2002). Differential regulation of HIV-1 clade-specific B, C, and E long terminal repeats by NF-kappaB and the Tat transactivator. *Virology* **296,** 77-83.

Ross, E.K., Buckler-White, A.J., Rabson, A.B., Englund, G. & Martin, M.A. (1991). Contribution of NF-kappa B and Sp1 binding motifs to the replicative capacity of human immunodeficiency virus type 1: distinct patterns of viral growth are determined by T-cell types. *J. Virol.* **65,** 4350-4358.

Rousseau, C., Abrams, E., Lee, M., Urbano, R. & King, M. (1997). long terminal repeat and *nef* gene variants of human immunodeficiency virus type 1 in perinatally infected long-term survivors and rapid progressors. *AIDS Res. Hum. Retroviruses.* **13,** 1611-1623.

Rowland-Jones, S., Sutton, J., Ariyoshi, K., Dong, T., Gotch, F., McAdam, S., Whitby, D., Sabally, S., Gallimore, A., Corrah, T., et al. (1995). HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women. *Nat. Med.* **1,** 59-64.

Roy, A.L., Meisterernst, M., Pognonec, P. & Roeder, R.G. (1991). Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature* **354,** 245-248.

Rubsamen-Waigmann, H., Briesen, H.V., Maniar, J.K., Rao, P.K., Scholz, C. & Pfutzner, A. (1991). Spread of HIV-2 in India [letter]. *Lancet* **337,** 550-551.

Ryu, S.E., Kwong, P.D., Truneh, A., Porter, T.G., Arthos, J., Rosenberg, M., Dai, X.P., Xuong, N.H., Axel, R., Sweet, R.W., et al. (1990). Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature* **348,** 419-426.

Sadaie, M.R., Rappaport, J., Benter, T., Josephs, S.F., Willis, R. & Wong-Staal, F. (1988). Missense mutations in an infectious human immunodeficiency viral genome: functional mapping of tat and identification of the rev splice acceptor. *Proc. Natl. Acad. Sci. U S A* **85,** 9224-9228.

Saksela, K. (1995). HIV-1 RNA in blood and pathogenesis of HIV infection. *Ann. Med.* **27,** 625-628.

Saltarelli, M., Querat, G., Konings, D.A., Vigne, R. & Clements, J.E. (1990). Nucleotide sequence and transcriptional analysis of molecular clones of CAEV which generate infectious virus. *Virology* **179,** 347-364.

Salvi, R., Garbuglia, A.R., Di Caro, A., Pulciani, S., Montella, F. & Benedetto, A. (1998). Grossly defective nef gene sequences in a human immunodeficiency virus type 1-seropositive long-term nonprogressor. *J. Virol.* **72,** 3646-3657.

Samb, N.D., Seck, K., Samb, B, et al. (1997). Cervico-vaginal shedding of HIV and STD among prostitutes and AIDS patients in Senegal. *Xth International Conference on AIDS and STD in Africa, Abidjan. [Abstract B.178].*

Santiago, M.L., Rodenburg, C.M., Kamenya, S., Bibollet-Ruche, F., Gao, F., Bailes, E., Meleth, S., Soong, S.J., Kilby, J.M., Moldoveanu, Z., Fahey, B., Muller, M.N., Ayouba, A., Nerrienet, E., McClure, H.M., Heeney, J.L., Pusey, A.E., Collins, D.A., Boesch, C., Wrangham, R.W., Goodall, J., Sharp, P.M., Shaw, G.M. & Hahn, B.H. (2002). SIVcpz in wild chimpanzees. *Science* **295,** 465.

Santiago, M.L., Lukasik, M., Kamenya, S., Li, Y., Bibollet-Ruche, F., Bailes, E., Muller, M.N., Emery, M., Goldenberg, D.A., Lwanga, J.S., Ayouba, A., Nerrienet, E., McClure, H.M., Heeney, J.L., Watts, D.P., Pusey, A.E., Collins, D.A., Wrangham, R.W., Goodall, J., Brookfield, J.F., Sharp, P.M., Shaw, G.M. & Hahn, B.H. (2003). Foci of endemic simian immunodeficiency virus infection in wild-living eastern chimpanzees (Pan troglodytes schweinfurthii). *J. Virol.* **77,** 7545-7562.

Sarr, A.D., Popper, S., Thior, I., Hamel, D.J., Sankale, J.L., Siby, T., Marlink, R., Essex, M., Mboup, S. & Kanki, P. (1999). Relation between HIV-2 proviral load and CD4+ lymphocyte count differs in monotypic and dual HIV infections. *J. Hum. Virol.* **2,** 45-51.

Schim van der Loeff, M.F. & Aaby, P. (1999). Towards a better understanding of the epidemiology of HIV-2. *AIDS* **13,** S69-84.

Schim van der Loeff, M.F., Jaffar, S., Aveika, A.A., Sabally, S., Corrah, T., Harding, E., Alabi, A., Bayang, A., Ariyoshi, K. & Whittle, H.C. (2002). Mortality of HIV-1, HIV-2 and HIV-1/HIV-2 dually infected patients in a clinic-based cohort in The Gambia. *A.I.D.S.* **16,** 1775-1783.

Schultz, T.F., Whitby, D., Hoad, J.G., Corrah, T., Whittle, H. & Weiss, R.A. (1990). Biological and molecular variability of human immunodeficiency virus type 2 isolates from The Gambia. *J. Virol.* **64,** 5177-5182.

Scriba, T.J., de Villiers, T., Treurnicht, F.K., zur Megede, J., Barnett, S.W., Engelbrecht, S. & van Rensburg, E.J. (2002). Characterization of the South African HIV type 1 subtype C complete 5' long terminal repeat, nef, and regulatory genes. *AIDS Res. Hum. Retroviruses* **18,** 149-159.

Sekigawa, I., Kaneko, H., Neoh, L.P., Takeda-Hirokawa, N., Akimoto, H., Hishikawa, T., Hashimoto, H., Hirose, S., Yamamoto, N. & Kaneko, Y. (1998). Differences of HIV envelope protein between HIV-1 and HIV-2: possible relation to the lower virulence of HIV-2. *Viral Immunol.* **11,** 1-8.

Sigurdsson, B., Palsson, P. & Grimsson, H. (1957). Visna, a demyelinating transmissible disease of sheep. *J. Neuropathol. Exp. Neurol.* **16,** 389-403.

Simm, M., Chao, W., Pekarskaya, O., Sova, P., Gupta, P., Balachandran, R. & Volsky, D.J. (1996). Genetic variability and function of the long terminal repeat from syncytium-inducing and non-syncytium-inducing human immunodeficiency virus type 1. *AIDS Res. Hum. Retroviruses* **12,** 801-809.

Simmons, G., Reeves, J.D., Hibbitts, S., Stine, J.T., Gray, P.W., Proudfoot, A.E. & Clapham, P.R. (2000). Co-receptor use by HIV and inhibition of HIV infection by chemokine receptor ligands. *Immunol. Rev.* **177,** 112-126.

Simmonds, P., Balfe, P., Peutherer, J.F., Ludlam, C.A., Bishop, J.O. & Brown, A.J. (1990). Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. *J. Virol.* **64,** 864-872.

Simon, F., Matheron, S., Tamalet, C., Loussert Ajaka, I., Bartczak, S., Pepin, J.M., Dhiver, C., Gamba, E., Elbim, C., Gastaut, J.A. & et al. (1993). Cellular and plasma viral load in patients infected with HIV-2. *AIDS* **7,** 1411-1417.

Smale, S.T. & Baltimore, D. (1989). The "initiator" as a transcription control element. *Cell* **57,** 103-113.

Smallman-Raynor, M. & Cliff, A. (1991). The spread of human immunodeficiency virus type 2 into Europe: a geographical analysis. *Int. J. Epidemiol.* **20,** 480-489.

Smith, N.A., Shaw, T., Berry, N., Vella, C., Okorafor, L., Taylor, D., Ainsworth, J., Choudhury, A., Daniels, R.S., El-Gadi, S., Fakoya, A., Moyle, G., Oxford, J., Tedder, R., O'Shea, S., de Ruiter, A. & Breuer, J. (2001). Antiretroviral therapy for HIV-2 infected patients. *J. Infect.* **42,** 126-133.

Soriano, V., Gutierrez, M., Caballero, E., Cilla, G., Fernandez, J.L., Aguilera, A., Tuset, C., Dronda, F., Martin, A.M., Carballo, E., Lopez, I. & Gonzalez-Lahoz, J. (1996). Epidemiology of HIV-2 infection in Spain. The HIV-2 Spanish Study Group. *Eur. J. Clin. Microbiol. Infect. Dis.* **15,** 383-388.

Soriano, V., Gomes, P., Heneine, W., Holguin, A., Doruana, M., Antunes, R., Mansinho, K., Switzer, W.M., Araujo, C., Shanmugam, V., Lourenco, H., Gonzalez-Lahoz, J. & Antunes, F. (2000). Human immunodeficiency virus type 2 (HIV-2) in Portugal: clinical spectrum , circulating subtypes, virus isolation, and plasma viral load. *J. Med. Virol.* **61,** 111-116.

Sousa, A.E., Chaves, A.F., Loureiro, A. & Victorino, R.M.M. (2001). Comparison of the frequency of interleukin (IL)-2-, Interferon-γ-, and IL-4- producing T cells in 2 diseases, human immunodeficiency virus types 1 and 2, with distinct clinical outcomes. *J.Infect.Dis.* **184,** 552-559.

Steel, C.M., Ludlam, C.A., Beatson, D., Peutherer, J.F., Cuthbert, R.J., Simmonds, P., Morrison, H. & Jones, M. (1988). HLA haplotype A1 B8 DR3 as a risk factor for HIV-related disease. *Lancet* **1,** 1185-1188.

Stein, D.S., Lyles, R.H., Graham, N.M., Tassoni, C.J., Margolick, J.B., Phair, J.P., Rinaldo, C., Detels, R., Saah, A. & Bilello, J. (1997). Predicting clinical progression or death in subjects with early-stage human immunodeficiency virus (HIV) infection: a comparative analysis of quantification of HIV RNA, soluble tumor necrosis factor type II receptors, neopterin, and beta2-microglobulin. Multicenter AIDS Cohort Study. *J. Infect. Dis.* **176,** 1161-1167.

Stewart, C.L., Schuetze, S., Vanek, M. & Wagner, E.F. (1987). Expression of retroviral vectors in transgenic mice obtained by embryo infection. *E.M.B.O. J.* **6,** 383-388.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12,** 599-606.

Sun, W., Graves, B.J. & Speck, N.A. (1995). Transactivation of the Moloney murine leukemia virus and T-cell receptor beta-chain enhancers by cbf and ets requires intact binding sites for both proteins. *J. Virol.* **69,** 4941-4949.

Swanstrom, R. & Wills, J.W. (1997). Retroviruses. *Cold Spring Harbour Laboratory press, Plainview, New York.* 263-334.

Switzer, W.M., Wiktor, S., Soriano, V., Silva-Graca, A., Mansinho, K., Coulibaly, I.M., Ekpini, E., Greenberg, A.E., Folks, T.M. & Heneine, W. (1998). Evidence of Nef truncation in human immunodeficiency virus type 2 infection. *J.Infect.Dis.* **177,** 65-71.

Tamalet, C., Simon, F., Dhiver, C., Matheron, S., de Micco, P., Gastao, J.A. & Brun-Vezinet, F. (1995). Autologous neutralizing antibodies and viral load in HIV-2-infected individuals. *AIDS* **9,** 90-91.

Tanaka, A., Salem, M., Eckroade, R.J. & Fujita, D.J. (1990). Characterization of avian retroviruses carrying activated transforming human c-src genes and of steps involved in expression of activated src-PKases in vitro. *Oncogene Res.* **5,** 305-22.

Teich, N. (1985). Taxonomy of Retroviruses. *RNA tumor viruses.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1-16.

Teleki, G. Understanding Chimpanzees. (1989). *Harvard Univ. Press, London.* 312-353.

Temin, H.M. & Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226,** 1211-1213.

Tersmette, M., van Dongen, J.J., Clapham, P.R., de Goede, R.E., Wolvers-Tettero, I.L., Geurts van Kessel, A., Huisman, J.G., Weiss, R.A. & Miedema, F. (1989). Human immunodeficiency virus infection studied in CD4-expressing human-murine T-cell hybrids. *Virology* **168,** 267-273.

Tevi-Benissan, C., Okome, M., Makuwa, M., Nkoume, M.N., Lansoud-Soukate, J., Georges, A., Georges-Courbot, M.C. & Belec, L. (1998). HIV-2 infection and HIV-1/HIV-2 dual reactivity in patients with and without AIDS-related symptoms in Gabon. *Emerg. Infect. Dis.* **4,** 130-131.

Thompson, C.C., Brown, T.A. & McKnight, S.L. (1991). Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. *Science* **253,** 762-768.

Thompson, C.B., Wang, C.Y., Ho, I.C., Bohjanen, P.R., Petryniak, B., June, C.H., Miesfeldt, S., Zhang, L., Nabel, G.J., Karpinski, B., et al. (1992). cis-acting sequences required for inducible interleukin-2 enhancer function bind a novel Ets-related protein, Elf-1. *Mol. Cell. Biol.* **12,** 1043-1053.

Tindall, B., Barker, S., Donovan, B., Barnes, T., Roberts, J., Kronenberg, C., Gold, J., Penny, R. & Cooper, D. (1988). Characterization of the acute clinical illness associated with human immunodeficiency virus infection. *Arch. Intern. Med.* **148,** 945-949.

Tirode, F., Busso, D., Coin, F. & Egly, J.M. (1999). Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol. Cell* **3,** 87-95.

Tobiume, M., Takahoko, M., Yamada, T., Tatsumi, M., Iwamoto, A. & Matsuda, M. (2002). Inefficient enhancement of viral infectivity and CD4 downregulation by human immunodeficiency virus type 1 Nef from Japanese long-term nonprogressors. *J. Virol.* **76,** 5959-5965.

Todd, J., Balira, R., Grosskurth, H., Mayaud, P., Mosha, F., ka-Gina, G., Klokke, A., Gabone, R., Gavyole, A., Mabey, D. & Hayes, R. (1997). HIV-associated adult mortality in a rural Tanzanian population. *A.I.D.S.* **11,** 801-807.

Tomiyama, H., Akari, H., Adachi, A. & Takiguchi, M. (2002). Different effects of Nef-mediated HLA class I down-regulation on human immunodeficiency virus type 1-specific CD8(+) T-cell cytolytic activity and cytokine production. *J. Virol.* **76,** 7535-7543.

Tong-Starksen, S.E., Welsh, T.M. & Peterlin, B.M. (1990). Differences in transcriptional enhancers of HIV-1 and HIV-2: response to T-cell activation signals. *J.Immunol.* **145,** 4348-4354.

Tong-Starksen, S.E., Baur, A., Lu, X.B., Peck, E. & Peterlin, B.M. (1993). Second exon of Tat of HIV-2 is required for optimal trans-activation of HIV-1 and HIV-2 LTRs. *Virology* **195,** 826-830.

Tristem, M., Mansinho, K., Champalimaud, J.L., Ayres, L. & Karpas, A. (1989). Six new isolates of human immunodeficiency virus type 2 (HIV-2) and the molecular characterization of one (HIV-2CAM2). *J. Gen. Virol.* **70,** 479-484.

Trono, D. (1995). HIV accessory proteins: leading roles for the supporting cast. *Cell* **82,** 189-192.

Tschochner, H., Sayre, M.H., Flanagan, P.M., Feaver, W.J. & Kornberg, R.D. (1992). Yeast RNA polymerase II initiation factor e: isolation and identification as the functional counterpart of human transcription factor IIB. *Proc. Natl. Acad. Sci. USA.* **89,** 11292-11296.

UNAIDS & WHO (2000). Report on the global HIV/AIDS epidemic.

Unutmaz, D., KewalRamani, V.N. & Littman, D.R. (1998). G protein-coupled receptors in HIV and SIV entry: new perspectives on lentivirus-host interactions and on the utility of animal models. *Semin. Immunol.* **10,** 225-236.

Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D. & Aloni, Y. (1992). Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell* **69,** 871-881.

van Baalen, C.A., Pontesilli, O., Huisman, R.C., Geretti, A.M., Klein, M.R., de Wolf, F., Miedema, F., Gruters, R.A. & Osterhaus, A.D. (1997). Human immunodeficiency virus type 1 Rev- and Tat-specific cytotoxic T lymphocyte frequencies inversely correlate with rapid progression to AIDS. *J. Gen. Virol.* **78,** 1913-1918.

Van der Ende, M.E., Schutten, M., Ly, T.D., Gruters, R.A. & Osterhaus, A.D. (1996). HIV-2 infection in 12 European residents: virus characteristics and disease progression. *AIDS* **10,** 1649-1655.

Van der Ende, M.E., Guillon, C., Boers, P.H., Ly, T.D., Gruters, R.A., Osterhaus, A.D. & Schutten, M. (2000). Antiviral resistance of biologic HIV-2 clones obtained from individuals on nucleoside reverse transcriptase inhibitor therapy. *J. Acquir. Immune Defic. Syndr.* **25,** 11-18.

Van der Maaten, M.J., Boothe, A.D. & Seger, C.L. (1972). Isolation of a virus from cattle with persistent lymphocytosis. *J. Natl. Cancer. Inst.* **49,** 1649-1657.

Van Lint, C., Amella, C.A., Emiliani, S., John, M., Jie, T. & Verdin, E. (1997). Transcription factor binding sites downstream of the human immunodeficiency virus type 1 transcription start site are important for virus infectivity. *J. Virol.* **71,** 6113-6127.

Vanden Haesevelde, M., Decourt, J.L., De Leys, R.J., Vanderborght, B., Van der Groen, G., Van Heuverswijn, H. & Saman, E. (1994). Genomic cloning and complete sequence analysis of a highly divergent African human immunodeficiency virus isolate. *J. Virol.* **68,** 1586-1596.

Vanden Haesevelde, M.M., Peeters, M., Jannes, G., Janssens, W., Van der Groen, G., Sharp, P.M. & Saman, E. (1996). Sequence analysis of a highly divergent HIV-1-related lentivirus isolated from a wild captured chimpanzee. *Virology* **221,** 346-350.

Varmus, H. (1988). Retroviruses. *Science,* 1427-1435.

Verhoef, K., Bauer, M., Meyerhans, A. & Berkhout, B. (1998). On the role of the second coding exon of the HIV-1 Tat protein in virus replication and MHC class I downregulation. *AIDS Res. Hum. Retroviruses* **14,** 1553-1559.

Verhoef, K., Sanders, R.W., Fontaine, V., Kitajima, S., Berkhout, B. (1999). Evolution of the human immunodeficiency virus type 1 long terminal repeat promoter by conversion of an NF-κB enhancer element into a GABP binding site. *J. Virol.* **73,** 1331-1340.

Viscidi, R.P., Mayur, K., Lederman, H.M. & Frankel, A.D. (1989). Inhibition of antigen-induced lymphocyte proliferation by Tat protein from HIV-1. *Science* **246,** 1606-1608.

Visco-Comandini, U., Yun, Z., Vahlne, A. & Sönnerborg, A. (1999). No association of HIV type 1 long terminal repeat sequence pattern with long-term nonprogression and *in vivo* viral replication levels in European subjects. *AIDS Res. Hum. Retroviruses.* **15,** 609-617.

Wagner, E.F., Vanek, M. & Vennstrom, B. (1985). Transfer of genes into embryonal carcinoma cells by retrovirus infection: efficient expression from an internal promoter. *E.M.B.O. J.* **4,** 663-666.

Wagner, H.U., Kamali, A., Nunn, A.J., Kengeya-Kayondo, J.F. & Mulder, D.W. (1993). General and HIV-1-associated morbidity in a rural Ugandan community. *AIDS* **7,** 1461-1467.

Wang, C.Y., Petryniak, B., Ho, I.C., Thompson, C.B. & Leiden, J.M. (1992). Evolutionarily conserved Ets family members display distinct DNA binding specificities. *J. Exp. Med.* **175,** 1391-1399.

Wang, C.Y., Bassuk, A.G., Boise, L.H., Thompson, C.B., Bravo, R. & Leiden, J.M. (1994). Activation of the granulocyte-macrophage colony-stimulating factor promoter in T cells requires cooperative binding of Elf-1 and AP-1 transcription factors. *Mol. Cell. Biol.* **14,** 1153-1159.

Wang, J.H., Yan, Y.W., Garrett, T.P., Liu, J.H., Rodgers, D.W., Garlick, R.L., Tarr, G.E., Husain, Y., Reinherz, E.L. & Harrison, S.C. (1990). Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature* **348,** 411-418.

Wawer, M.J., Serwadda, D., Musgrave, S.D., Konde-Lule, J.K., Musagara, M. & Sewankambo, N.K. (1991). Dynamics of spread of HIV-I infection in a rural district of Uganda. *B.M.J.* **303,** 1303-1306.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H. & Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92,** 451-462.

Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., et al. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 1995 **373,** 117-122.

Weinzierl, R.O., Dynlacht, B.D. & Tjian, R. (1993). Largest subunit of Drosophila transcription factor IID directs assembly of a complex containing TBP and a coactivator. *Nature* **362,** 511-517.

Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J. & Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387,** 426-430.

Werner, J. & Gledерblom, H. (1979). Isolation of a foamy virus from patients with de Quervain thyroiditis. *Lancet* **2,** 258-259.

West, M.J., Lowe, A.D. & Karn, J. (2001). Activation of human immunodeficiency virus transcription in T cells revisited: NF-kappaB p65 stimulates transcriptional elongation. *J. Virol.* **75,** 8524-8537.

Westby, M., Manca, F. & Dalgleish, A.G. (1996). The role of host immune responses in determining the outcome of HIV infection. *Immunol. Today* **17,** 120-126.

Whittle, H., Morris, J., Todd, J., Corrah, T., Sabally, S., Bangali, J., Ngom, P.T., Rolfe, M. & Wilkins, A. (1994). HIV-2-infected patients survive longer than HIV-1-infected patients. *AIDS* **8,** 1617-1620.

Whittle, H., Yamuah, L.K., O'Donovan, D., Milligan, P., Harding, E. & Wilkins H.A. (1998a). HIV-1 and HIV-2 infection in pregnant women in The Gambia : risk factors. *12th World AIDS Conference, Geneva. [Abstract 23164].*

Whittle, H.C., Ariyoshi, K. & Rowland-Jones, S. (1998b). HIV-2 and T cell recognition. *Curr. Opin. Immunol.* **10,** 382-387.

Wild, C., Dubay, J.W., Greenwell, T., Baird, T. Jr., Oas, T.G., McDanal, C., Hunter, E. & Matthews, T. (1994). Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex. *Proc. Natl. Acad. Sci. USA.* **91,** 12676-12680.

Wilkins, A., Oelman, B., Pepin, J., Cham, K., Corrah, T., Hughes, A., Manneh, K., Manneh, K., Njai, R., Rolfe, M., et al. (1991a). Trends in HIV-1 and HIV-2 infection in The Gambia. *AIDS* **5,** 1529-1530.

Wilkins, A., Hayes, R., Alonso, P., Baldeh, S., Berry, N., Cham, K., Hughes, A., Jaiteh, K., Oelman, B., Tedder, R., et al. (1991b). Risk factors for HIV-2 infection in The Gambia. *AIDS* **5,** 1127-1132.

Wilkins, A., Ricard, D., Todd, J., Whittle, H., Dias, F. & Paulo Da Silva, A. (1993). The epidemiology of HIV infection in a rural area of Guinea-Bissau. *AIDS* **7,** 1119-1122.

Wimmer, J., Fujinaga, K., Taube, R., Cujec, T.P., Zhu, Y., Peng, J., Price, D.H. & Peterlin, B.M. (1999). Interactions between Tat and TAR and human immunodeficiency virus replication are facilitated by human cyclin T1 but not cyclins T2a or T2b. *Virology* **255,** 182-189.

Winkler, C., Modi, W., Smith, M.W., Nelson, G.W., Wu, X., Carrington, M., Dean, M., Honjo, T., Tashiro, K., Yabe, D., Buchbinder, S., Vittinghoff, E., Goedert, J.J., O'Brien, T.R., Jacobson, L.P., Detels, R., Donfield, S., Willoughby, A., Gomperts, E., Vlahov, D., Phair, J. & O'Brien, S.J. (1998). Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC). *Science* **279,** 389-393.

Witte, O.N., Rosenberg, N., Paskind, M., Shields, A. & Baltimore D. (1978). Identification of an Abelson murine leukemia virus-encoded protein present in transformed fibroblast and lymphoid cells. *Proc. Natl. Acad. Sci. USA.* **75,** 2488-2492.

Wolf, D., Witte, V., Laffert, B., Blume, K., Stromer, E., Trapp, S., d'Aloja, P., Schurmann, A. & Baur, A.S. (2001). HIV-1 Nef associated PAK and PI3-kinases stimulate Akt-independent Bad-phosphorylation to induce anti-apoptotic signals. *Nat. Med.* **7,** 1217-1224.

Wolinsky, S.M., Korber, B.T., Neumann, A.U., Daniels, M., Kunstman, K.J., Whetsell, A.J., Furtado, M.R., Cao, Y., Ho, D.D. & Safrit, J.T. (1996). Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* **272,** 537-542.

Wu, L., Gerard, N.P., Wyatt, R., Choe, H., Parolin, C., Ruffing, N., Borsetti, A., Cardoso, A.A., Desjardin, E., Newman, W., Gerard, C. & Sodroski, J. (1996). CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* **384,** 179-183.

Wu-Baer, F., Lane, W.S. & Gaynor, R.B. (1998). Role of the human homolog of the yeast transcription factor SPT5 in HIV-1 Tat-activation. *J. Mol. Biol.* **277,** 179-197.

Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A. & Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393,** 705-711.

Xu, X.N., Screaton, G.R., Gotch, F.M., Dong, T., Tan, R., Almond, N., Walker, B., Stebbings, R., Kent, K., Nagata, S., Stott, J.E. & McMichael, A.J. (1997). Evasion of cytotoxic T lymphocyte (CTL)

responses by nef-dependent induction of Fas ligand (CD95L) expression on simian immunodeficiency virus-infected cells. *J. Exp. Med.* **186,** 7-16.

Xu, X.N., Laffert, B., Screaton, G.R., Kraft, M., Wolf, D., Kolanus, W., Mongkolsapay, J., McMichael, A.J. & Baur, A.S. (1999). Induction of Fas ligand expression by HIV involves the interaction of Nef with the T cell receptor zeta chain. *J. Exp. Med.* **189,** 1489-1496.

Yamamoto, J.K., Sparger, E., Ho, E.W., Andersen, P.R., O'Connor, T.P., Mandell, C.P., Lowenstine, L., Munn, R. & Pedersen, N.C. (1988). Pathogenesis of experimentally induced feline immunodeficiency virus infection in cats. *Am. J. Vet. Res.* **49,** 1246-1258.

Yan, Q., Moreland, R.J., Conaway, J.W. & Conaway, R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J. Biol. Chem.* **274,** 35668-35675.

Yang, X., Herrmann, C.H. & Rice, A.P. (1996). The human immunodeficiency virus Tat proteins specifically associate with TAK in vivo and require the carboxyl-terminal domain of RNA polymerase II for function. *J. Virol.* **70,** 4576-4584.

Yang, X., Gold, M.O., Tang, D.N., Lewis, D.E., Aguilar-Cordova, E., Rice, A.P. & Herrmann, C.H. (1997). TAK, an HIV Tat-associated kinase, is a member of the cyclin-dependent family of protein kinases and is induced by activation of peripheral blood lymphocytes and differentiation of promonocytic cell lines. *Proc. Natl. Acad. Sci. USA.* 94, 12331-12336.

Yoshida, M., Miyoshi, I. & Hinuma, Y. (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci. USA.* **79,** 2031-2035.

Zagury, D., Lachgar, A., Chams, V., Fall, L.S., Bernard, J., Zagury, J.F., Bizzini, B., Gringeri, A., Santagostino, E., Rappaport, J., Feldman, M., Burny, A. & Gallo, R.C. (1998). Interferon alpha and Tat involvement in the immunosuppression of uninfected T cells and C-C chemokine decline in AIDS. *Proc. Natl. Acad. Sci. USA.* **95,** 3851-3856.

Zawel, L., Kumar, K.P. & Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* **9,** 1479-1490.

Zekeng, L., Obiang Sima, J., Hampl, H., Ndemesogo, J.M., Ntutumu, J., Sima, V., Devare, S., Kaptue, L. & Gurtler, L. (1997). Update on HIV-1 group O infection in Equatorial Guinea, Central Africa. *AIDS* **11,** 1410-1412.

Zenzie-Gregory, B., Sheridan, P., Jones, K.A. & Smale, S.T. (1993). HIV-1 core promoter lacks a simple initiator element but contains a bipartite activator at the transcription start site. *J. Biol. Chem.* **268,** 15823-15832.

Zhang, Y., Nakata, K., Weiden, M. & Rom, W.N. (1995). Mycobacterium tuberculosis enhances human immunodeficiency virus-1 replication by transcriptional activation at the long terminal repeat. *J. Clin. Invest.* **95,** 2324-2331.

Zhang, L., Huang, H., Yuan, B.K., Chen, J.I., & Ho, D.D. (1997a). Identification of a replication-competent pathogenic human immunodeficiency virus type 1 with a duplication in the TCF-region but lacking NF-κB sites. *J. Virol.* **71,** 1651-1656.

Zhang, L., Huang, H., Yuan, B.K., Chen, J.I., Ip, J., & Ho, D.D. (1997b). Genotypic and phenotypic characterisation of long terminal repeat sequences from long-term survivors of human immunodeficiency virus type 1 infection. *J. Virol.* **71,** 5608-5613.

Zhou, H. & Duesberg, P.H. (1990). A retroviral promoter is sufficient to convert proto-src to a transforming gene that is distinct from the src gene of Rous sarcoma virus. *Proc. Natl. Acad. Sci. U S A.* **87,** 9128-32.

Zhou, W., Parent, L.J., Wills, J.W. & Resh, M.D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J. Virol.* **68,** 2556-2569.

Zhou, Q., Chen, D., Pierstorff, E. & Luo, K. (1998). Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. *EMBO J.* **17,** 3681-3691.

Zhou, M., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J., Price, D.H. & Brady, J.N. (2000). Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol. Cell. Biol.* **20,** 5077-5086.

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M.B. & Price, D.H. (1997). Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev.* **11,** 2622-2632.

Appendix I.

Chemicals and Buffer formulae.

TE buffer:

10mM Tris-HCl (pH 7.6), 1mM EDTA

Phosphate buffered saline (PBS):

150mM NaCl, 15mM KH_2PO_4, 63mM $Na_2HPO_4xH_2O$, 26mM KCl, pH 7.2

PCR and cloning reagents:

10 x *Pfu* Turbo reaction buffer:

100mM KCl, 100mM $(NH_4)_2SO_4$, 200mM Tris-HCl (pH 8.8), 20mM $MgSO_4$, 1% Triton X-100 and 1mg/ml nuclease-free BSA

50 x Tris Acetate EDTA (TAE):

242g/L Tris base, 57.1ml/L glacial acetic acid, 37.2g/L EDTA, adjusted to pH 8.1

1% Agarose gel:

1 x TAE buffer containing 1.5 to 2% (w/v) agarose (Promega) and 0.5mg/ml ethidium bromide

Agarose gel loading buffer (10x):

40% sucrose, 0.1% Orange G in dH_2O

1kb DNA Marker (Life Technologies):

20µl 1kb marker: diluted in 200µl of TE and 80µl of loading dye

Buffer H:

50mM Tris-HCl, 10mM $MgCl_2$, 100mM NaCl, 1mM Dithioerythriol, pH 7.5

10 x T4 DNA ligase buffer:

200mM TrisCl (pH 7.6), 50mM $MgCl_2$, 50mM dithiothreitol, 500µg/ml bovine serum albumin, 0.5M ATP

DNA concentration assay diluent:

10mM Tris HCl, pH 7.5, containing 1mM EDTA, 100mM NaCl

Bacterial solutions:

2 x TY Broth:

16g/L Bacto-tryptone, 10g/L Bacto-yeast extract, NaCl 5g/L, pH7.0

2 x TY Agar:

2 x TY broth with 1.5% (w/v) Difco Bacto agar

Antibiotic Solution :

Ampicillin stock solution 100mg/ml in dH_2O stored at –20°C. Working stock concentration 50µg/ml

Calcium Chloride Bacterial Transformation Buffer:

100mM Calcium Chloride, 47g $CaCl_2.H_2O$, dissolved in 100ml of dH_2O and sterilised by autoclaving, stored at –20°C

Tissue culture media:

Complete medium DMEM:

Dulbecco's Modified Eagle's Medium supplemented with 10% FCS, 5mM L-glutamine, 100U/ml penicillin, 100µg/ml streptomycin, 0.075% sodium bicarbonate. Stored at 4°C.

Complete medium RPMI 1640:

RPMI 1640 supplemented with 10% FCS, 5mM L-glutamine, 100U/ml penicillin, 100µg/ml streptomycin, 0.075% sodium bicarbonate. Adjusted to pH 7.0 with NaOH. Stored at 4°C.

2 x Hepes Buffered Saline:

8.18g NaCl, 5.95g HEPES (tissue culture grade), 0.2g Na_2HPO_4, 400ml dH_2O adjusted to pH 7.1 with NaOH, topped up to 500ml with dH_2O. Filter sterilised and stored at 4°C.

Freezing Medium:

10% DMSO, 40% FCS, 50% DMEM/RPMI

Luciferase reporter assay reagents:

Reagents including Passive Lysis Buffer, Luciferase assay reagent II, and Stop & Glo™ all supplied by Promega. Composition undisclosed.

Sequencing reagents:

Thermo Sequenase pre-mixed cycle sequencing kit:

This kit contains pre-mixed A, C, G, and T reagents. Each reagent contains Thermo Sequenase DNA polymerase, and the corresponding dNTPs and ddNTPs. Exact reagent compositions are undisclosed (Amersham Life Science).

10 x Tris Borate EDTA (TBE):

110g/L Tris base, 55g/L boric acid, 9.3g/L EDTA, adjusted to pH 8.3

10% ammonium persulphate solution (APS):

1g APD dissolved in 10ml dH_2O

6% denaturing polyacrylamide/urea sequencing gel:

11g urea, 3ml bisacrylamide (contains 40% w/v acrylamide – 2.105% w/v bisacrylamide), 1.5ml 40% acrylamide (acrylamide:bisacrylamide ratio 19:1), 3ml of 10x TBE, 12.4ml dH_20, . Polymerised by the addition of 150µl of 10% ammonium persulphate solution (APS) and 15µl TEMED to 30ml of gel mix.

Sequencing gel tank buffer 1 x TBE:

60ml 10 x TBE in 540ml dH_2O

Denaturing loading dye:

90% (v/v) formamide

Isoelectric solution:

Loading dye/salt solution/dH_2O, in a ratio of 2:1:1

Appendix II.

Plasmid Maps.

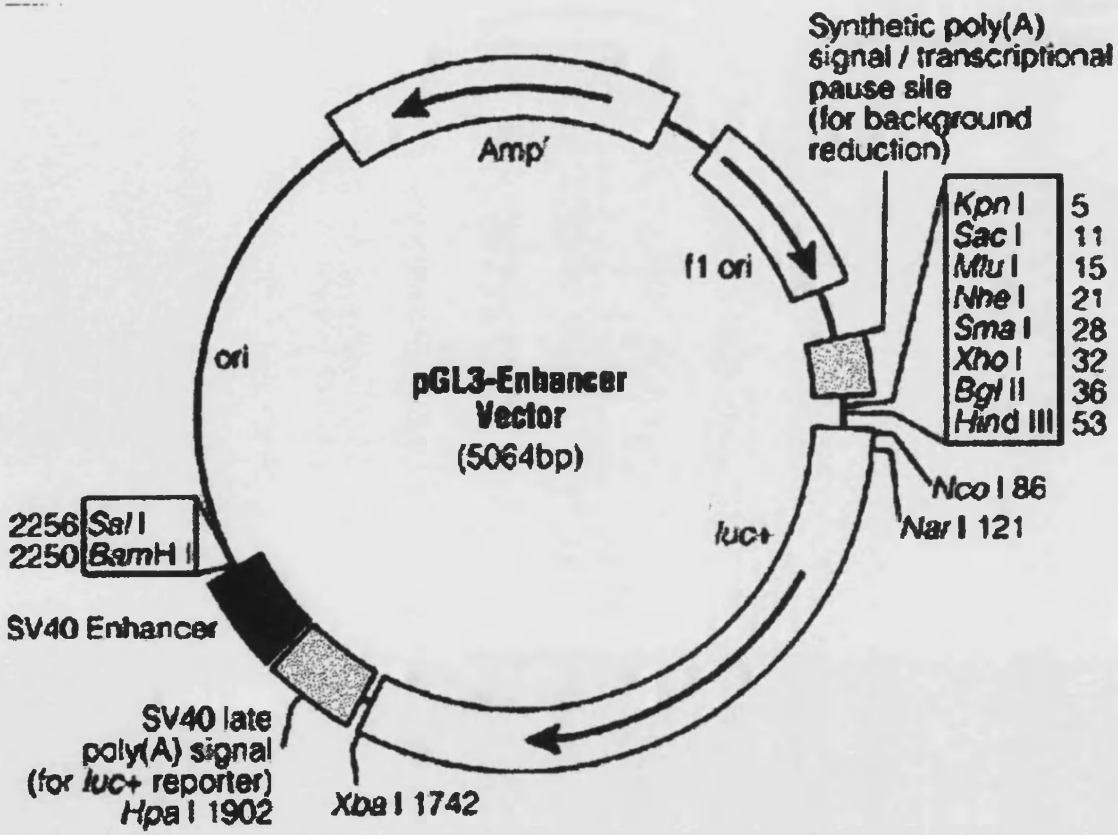


Figure II.1: Schematic map for pGL3-Enhancer (pGL3E) (Promega). This reporter vector contains the firefly luciferase (*Photinus pyralis*) gene. The polylinker region located upstream of the luciferase gene allowed the rapid and efficient insertion of HIV-2 LTR. The *Mlu I* and *Bbl II* sites were utilised in this study. Expression of firefly luciferase was under the direct control of the HIV-2 promoter. The SV40 enhancer located downstream of the firefly luciferase gene aided the verification of functional promoter elements inserted into pGL3E by increasing the level of firefly luciferase gene transcription. In addition, an ampicillin-resistance gene, β-lactamase, and a high copy number prokaryotic origin of replication allowed pGL3E selection, maintenance, and manipulation in *E.coli* hosts. The SV40 late poly(A) signal increased the efficiency of transcription, termination and polyadenylation of the luciferase transcripts.

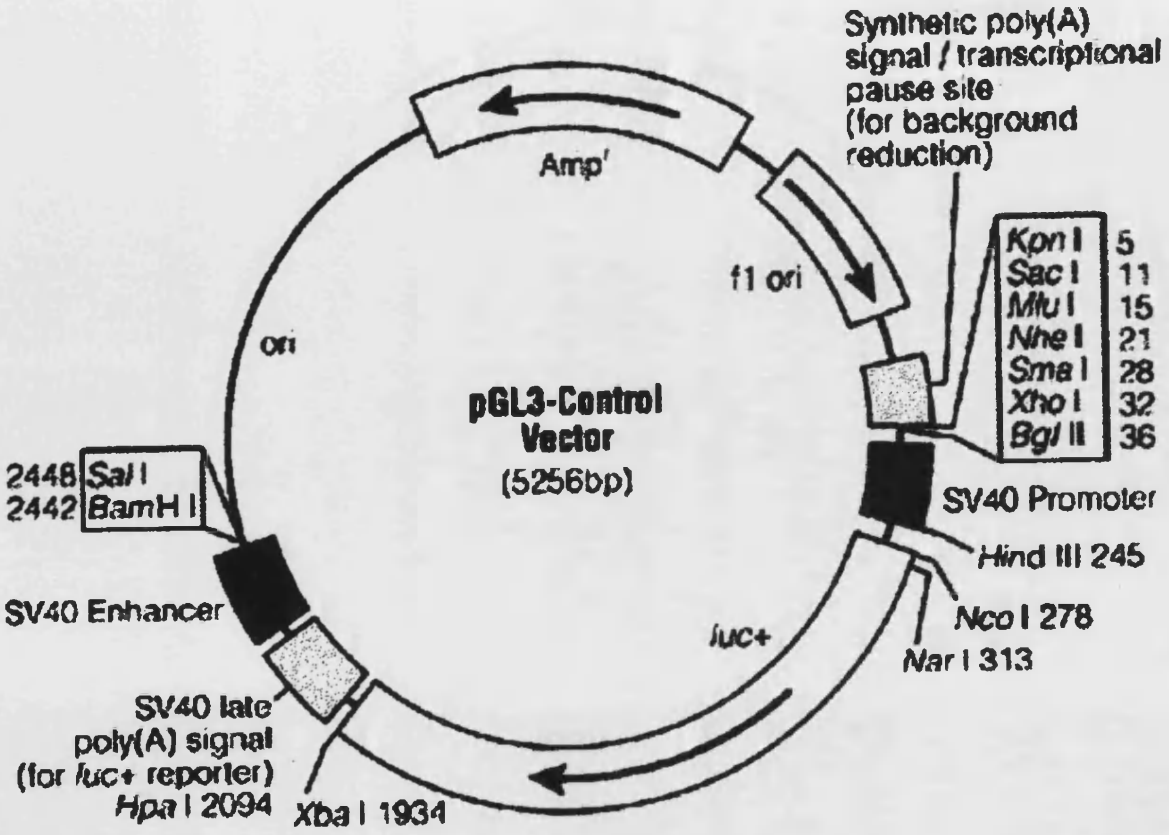


Figure II.2: Schematic map for pGL3-Control (pGL3C) (Promega). This reporter vector contains the firefly luciferase (*Photinus pyralis*) gene under the control of the SV40 promoter, resulting in strong expression of *luc+*. This reporter vector was used as a positive control for firefly luciferase expression in the experiments performed to determine the *trans* effects experienced by cotransfected vectors (Chapter 4).

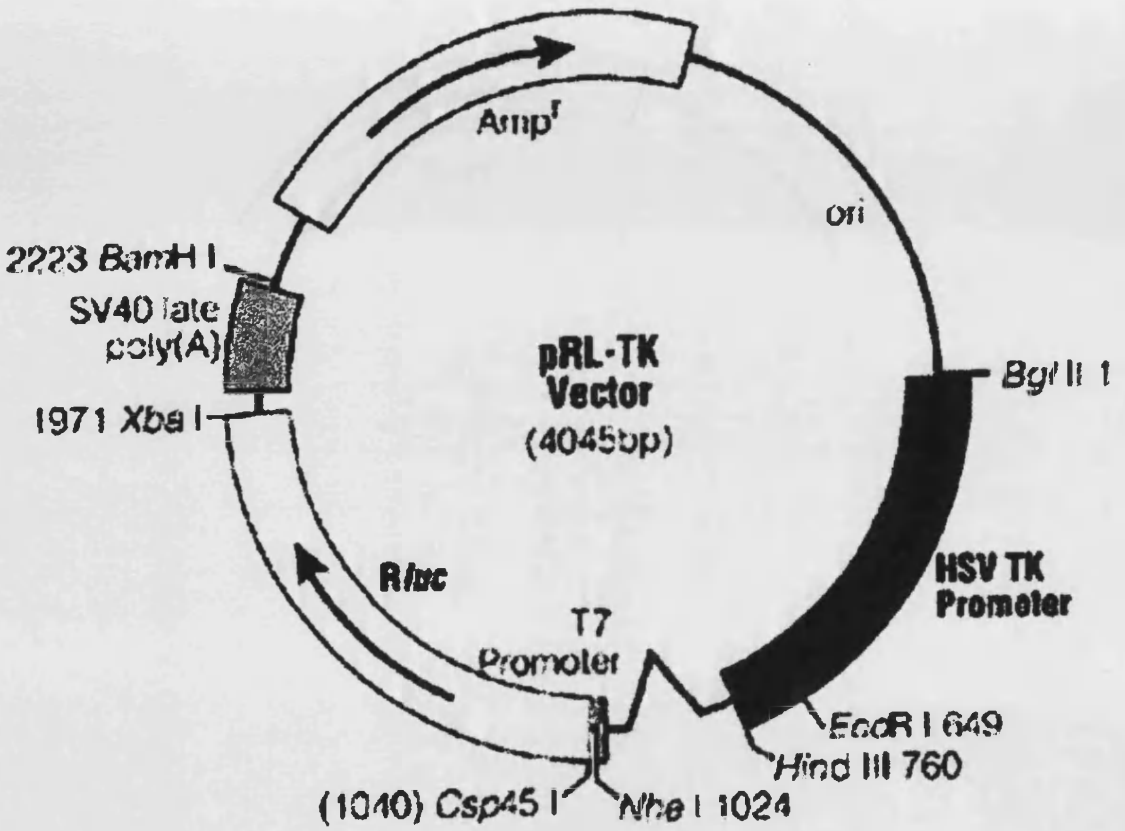


Figure II.3: Schematic map for pRL-TK (Promega). This internal control reporter vector contains the *Renilla* luciferase (*Renilla reniformis*) gene under the control of the herpes simplex virus thymidine kinase (HSV-TK) promoter, providing low to moderate levels of *Renilla* luciferase expression in cotransfected mammalian cells. This reporter vector was chosen for use as the internal control vector in cotransfections performed with pGL3E containing HIV-2 LTR inserts. (Chapter 4, 5, and 6). Cotransfections were performed at an experimental reporter vector:internal control vector ratio of 10:1.

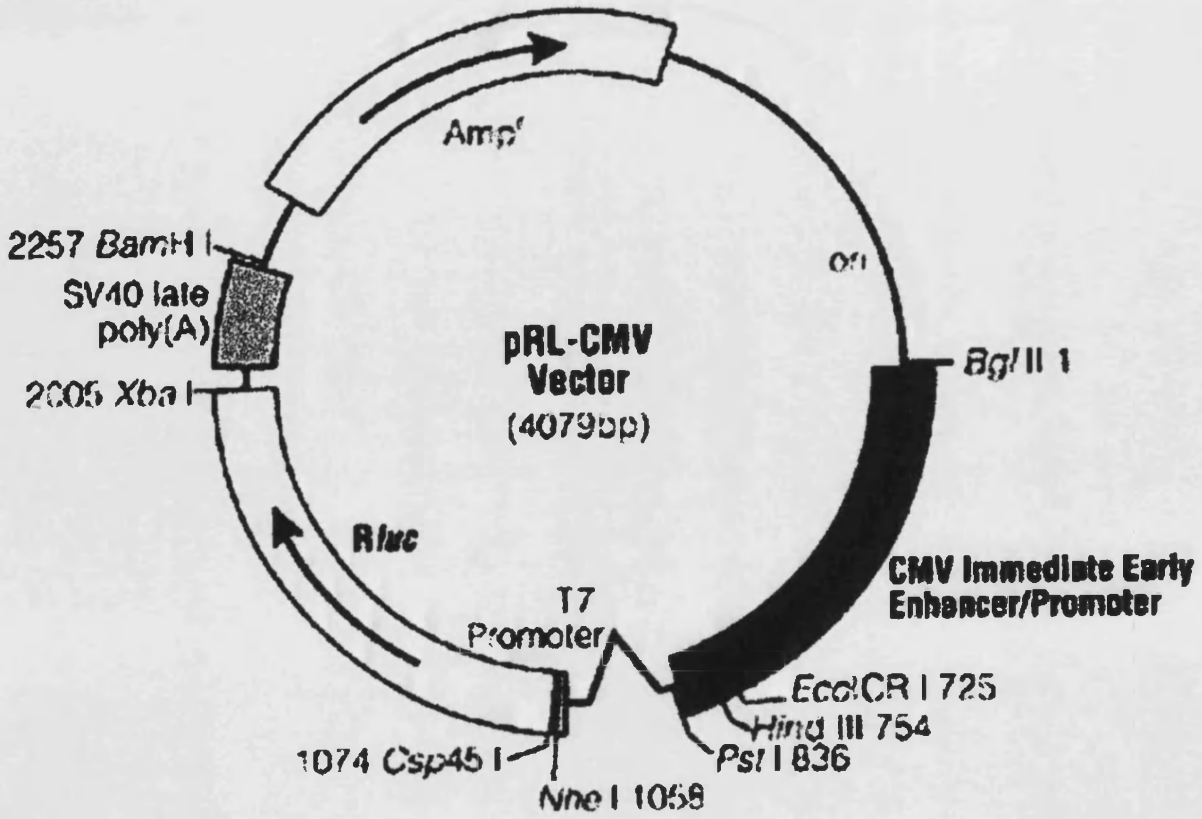


Figure II.4: Schematic map for pRL-CMV (Promega). This internal control reporter vector contains the *Renilla* luciferase (*Renilla reniformis*) gene under the control of the cytomegalovirus (CMV) promoter, providing high levels of *Renilla* luciferase expression in cotransfected mammalian cells. This reporter vector was used in the assessment of the optimal experimental reporter:internal control vector pairing (Chapter 4).

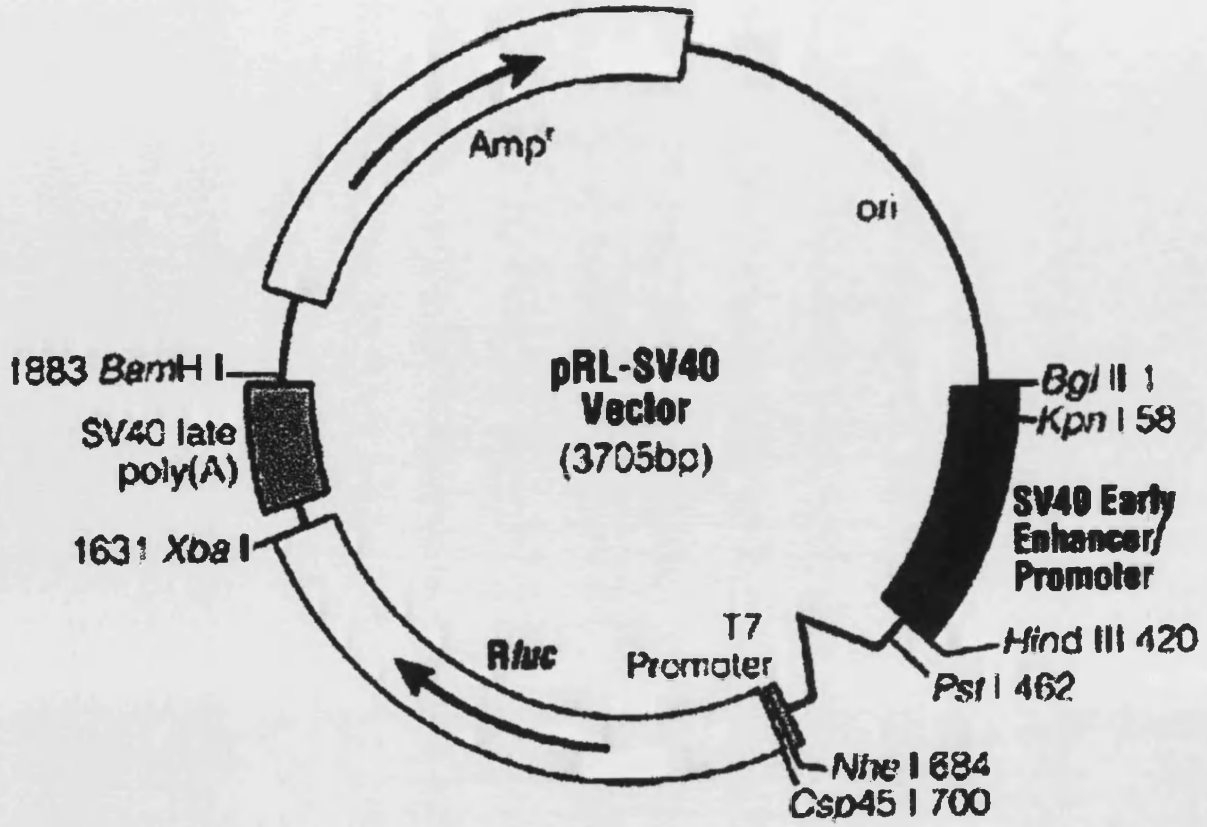


Figure II.5: Schematic map for pRL-SV40 (Promega). This internal control reporter vector contains the *Renilla* luciferase (*Renilla reniformis*) gene under the control of the SV40 promoter, providing high levels of *Renilla* luciferase expression in cotransfected mammalian cells. This reporter vector was used in the assessment of the optimal experimental reporter:internal control vector pairing (Chapter 4).

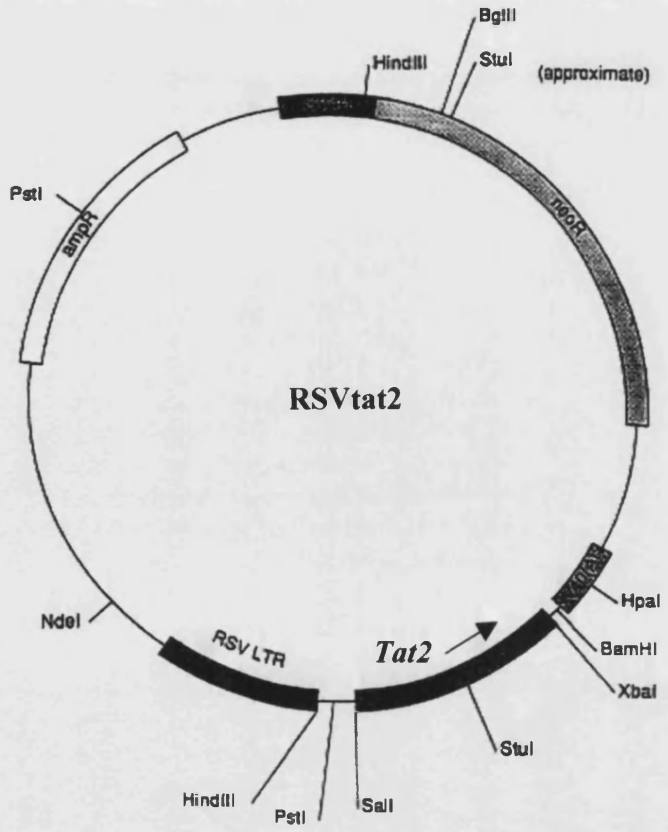


Figure II.6: Schematic map for RSVtat2 (kind gift from Dr.Browning, University of Michegan). This vector was used for the expression of HIV-2 Tat (ROD) when transfected into mammalian cell-lines. Expression of the first two exons of HIV-2 Tat (ROD) is under the control of the Rous Sarcoma Virus (RSV) promoter.